BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

VOLUME IX

JANUARY 1913 TO OCTOBER 1913

CAMBRIDGE
AT THE UNIVERSITY PRESS

LONDON: FETTER LANE, E.C.
(C. F. CLAY, MANAGER)
AND H. K. LEWIS, GOWER STREET
WILLIAM WESLEY AND SON, 28, ESSEX STREET, STRAND
EDINBURGH: 100, PRINCES STREET
CHICAGO: UNIVERSITY OF CHICAGO PRESS
BERLIN: A. ASHER AND CO.
LEIPSIC: BROCKHAUS
BOMBAY AND CALCUTTA: MACMILLAN AND CO., LIMITED
TORONTO: J. M. DENT AND SONS, LIMITED
TOKYO: THE MARUZEN-KABUSHIKI-KAISHA

Cambridge:
PRINTED BY JOHN CLAY, M.A.
AT THE UNIVERSITY PRESS.

CONTENTS OF VOL. IX.

Memoirs.

Miscellanea.

Plates.

Vol. IX. Parts I and II. March 1913

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

CAMBRIDGE
AT THE UNIVERSITY PRESS

LONDON: FETTER LANE, E.C.
(C. F. CLAY, MANAGER)
AND H. K. LEWIS, GOWER STREET
WILLIAM WESLEY AND SON, 28, ESSEX STREET, STRAND
EDINBURGH: 100, PRINCES STREET
CHICAGO: UNIVERSITY OF CHICAGO PRESS
BERLIN: A. ASHER AND CO.
LEIPSIC: BROCKHAUS
BOMBAY AND CALCUTTA: MACMILLAN AND CO., LIMITED

Entered at the New York Post Office as Second Class Matter

Price Twenty Shillings net.

[*Issued March* 12, 1913]

VOLUME IX MARCH, 1913 Nos. 1 AND 2

BIOMETRIKA.

ON THE PROBABLE ERRORS OF FREQUENCY CONSTANTS.

PART II.

EDITORIAL.

(1) THE probable errors of frequency constants for a single variable have been discussed in an elementary manner in Vol. II. pp. 273—281 of this Journal. The following notes on the probable errors of frequency constants of distributions with two variables have been provided at the request of certain of our readers*.

All the constants of a frequency distribution of two variates are expressible in terms of the higher product moments defined by

$$p_{q,q'} = S\{n_{ss'} x_s{}^q y_{s'}{}^{q'}\}/N \text{ ....(i)},$$

where the origin is at *fixed* values of x and y. Transferred to the mean $\bar{x}, \bar{y}$, we have:

$$\bar{p}_{q,q'} = S\{n_{ss'}(x_s - \bar{x})^q (y_{s'} - \bar{y})^{q'}\}/N \text{ ....(ii)},$$

where $\bar{x}$ and $\bar{y}$ will vary from sample to sample. Here $n_{ss'}$ is the frequency of individuals with characters x_s and $y_{s'}$, and N is the total population.

Throughout $r_{\alpha\beta}$ will denote the correlation between two quantities α and β; σ_α, σ_β will denote the standard deviations of α and β.

The following results are well known:

$$\sigma^2_{n_{ss'}} = n_{ss'}\left(1 - \frac{n_{ss'}}{N}\right) \text{ ....(iii)},$$

$$\sigma_{n_{ss'}}\sigma_{n_{tt'}} r_{n_{ss'}n_{tt'}} = -\frac{n_{ss'}n_{tt'}}{N} \text{ ....(iv)}.$$

* Reproduced from Lecture Notes.

Further if $n_s = S(n_{ss'})$ for all values of s',

$$\sigma_{n_s}\sigma_{n_{s'}} r_{n_s n_{s'}} = n_{ss'} - \frac{n_s n_{s'}}{N} \quad \text{(v)},$$

$$\sigma_{n_s}\sigma_{n_{ts'}} r_{n_s n_{ts'}} = -\frac{n_s n_{ts'}}{N} \quad \text{(vi)},$$

$$\sigma_{n_s}\sigma_{n_{ss'}} r_{n_s n_{ss'}} = n_{ss'}\left(1 - \frac{n_s}{N}\right) \quad \text{(vii)}.$$

See *Biometrika*, Vol. v. pp. 191—2.

(2) We have, if δ denote a variation in any frequency constant due to random sampling,

$$N\delta p_{q,q'} = S(\delta n_{ss'} x_s^q y_{s'}^{q'}) \quad \text{(vii) } bis,$$

and if m be number of random samples:

$$\begin{aligned} N^2\sigma^2_{p_{q,q'}} &= \frac{1}{m}\Sigma(\delta p_{q,q'})^2 \\ &= S(\sigma^2_{n_{ss'}} x_s^{2q} y_{s'}^{2q'}) + 2S(\sigma_{n_{ss'}}\sigma_{n_{tt'}} r_{n_{ss'}n_{tt'}} \times x_s^q x_t^q y_{s'}^{q'} y_{t'}^{q'}) \\ &= Sn_{ss'}\left(1 - \frac{n_{ss'}}{N}\right) x_s^{2q} y_{s'}^{2q'} - 2S\left(\frac{n_{ss'}n_{tt'}}{N} x_s^q x_t^q y_{s'}^{q'} y_{t'}^{q'}\right) \\ &= N(p_{2q,2q'} - p_{q,q'}p_{q,q'}). \end{aligned}$$

Thus
$$\sigma^2_{p_{q,q'}} = \frac{p_{2q,2q'} - p^2_{q,q'}}{N} \quad \text{(viii)}.$$

Again
$$N\delta p_{u,u'} = S(\delta n_{ss'} x_s^u y_{s'}^{u'}),$$

$$\begin{aligned} N^2\sigma_{p_{q,q'}}\sigma_{p_{u,u'}} r_{p_{q,q'}p_{u,u'}} &= S(\sigma_{n_{ss'}}\sigma_{n_{tt'}} r_{n_{ss'}n_{tt'}} \times x_s^{q+u} x_t y_{s'}^{q'} y_{t'}^{u'}) + S(\sigma^2_{n_{ss'}} x_s^{q+u} y_{s'}^{q'+u'}) \\ &= N(p_{q+u,q'+u'} - p_{q,q'}p_{u,u'}), \end{aligned}$$

thus
$$\sigma_{p_{q,q'}}\sigma_{p_{u,u'}} r_{p_{q,q'}p_{u,u'}} = \frac{p_{q+u,q'+u'} - p_{q,q'}p_{u,u'}}{N} \quad \text{(ix)}.$$

(viii) and (ix) refer only to the higher moment coefficients about a *fixed* origin.

About the mean we have

$$\begin{aligned} N\bar{p}_{q,q'} &= S\{n_{ss'}(x_s - \bar{x})^q (y_{s'} - \bar{y})^{q'}\}, \\ N\delta\bar{p}_{q,q'} &= S\{\delta n_{ss'}(x_s - \bar{x})^q (y_{s'} - \bar{y})^{q'}\} - q\delta\bar{x}\bar{p}_{q-1,q'} - q'\delta\bar{y}\bar{p}_{q,q'-1} \quad \text{(x)}. \end{aligned}$$

Now it is clear that before going further we want to know the correlation between variations in $n_{ss'}$ and $\bar{x}$ or $\bar{y}$. Now

$$N\bar{x} = S(n_t x_t),$$
$$N\delta\bar{x} = S(\delta n_t x_t),$$
$$N\delta\bar{x}\,\delta n_{ss'} = S(\delta n_{ss'}\delta n_t x_t),$$
$$\begin{aligned} N\sigma_{\bar{x}}\sigma_{n_{ss'}} r_{\bar{x}n_{ss'}} &= n_{ss'}\left(1 - \frac{n_s}{N}\right) x_s - S\left(\frac{n_t n_{ss'}}{N} x_t\right) \\ &= n_{ss'}(x_s - \bar{x}), \text{ using (vi) and (vii)}, \end{aligned}$$

or
$$\sigma_{\bar{x}}\sigma_{n_{ss'}} r_{\bar{x}n_{ss'}} = \frac{n_{ss'}}{N}(x_s - \bar{x}) \quad \text{(xi)}.$$

Next we want $\sigma_{\bar{x}}\sigma_{\bar{y}}r_{\bar{x}\bar{y}}$, but this is known, or easily found to be $\bar{p}_{11}/N$.

Returning to (x), squaring, summing for every possible sample and dividing by the number of samples, we have

$$N^2\sigma^2_{\bar{p}_{q,q'}} = \bar{p}_{2q,2q'} - \bar{p}^2_{q,q'} + q^2\bar{p}_{2,0}\bar{p}^2_{q-1,q'} + q'^2\bar{p}_{0,2}\bar{p}^2_{q,q'-1} + 2qq'\bar{p}_{1,1}\bar{p}_{q-1,q'} \times \bar{p}_{q,q'-1}$$
$$- 2q\bar{p}_{q+1,q'}\bar{p}_{q-1,q'} - 2q'\bar{p}_{q,q'+1}\bar{p}_{q,q'-1} \quad \text{...(xii)}.$$

We may now find the correlation between any pair of higher moment coefficients,

$$N\delta\bar{p}_{u,u'} = S\{\delta n_{ss'}(x_s - \bar{x})^u(y_{s'} - \bar{y})^{u'}\} - u\delta\bar{x}\bar{p}_{u-1,u'} - u'\delta\bar{y}\bar{p}_{u,u'-1} \quad \text{...(xiii)}.$$

Multiplying (x) and (xiii) together, summing for all possible samples and dividing by the number of samples and N, we have

$$N\sigma_{\bar{p}_{q,q'}}\sigma_{\bar{p}_{u,u'}}r_{\bar{p}_{q,q'}\bar{p}_{u,u'}} = \bar{p}_{q+u,q'+u'} - \bar{p}_{q,q'}\bar{p}_{u,u'} + qu\bar{p}_{2,0}\bar{p}_{q-1,q'}\bar{p}_{u-1,u'} + q'u'\bar{p}_{0,2}\bar{p}_{q,q'-1}\bar{p}_{u,u'-1}$$
$$+ qu'\bar{p}_{1,1}\bar{p}_{q-1,q'}\bar{p}_{u,u'-1} + q'u\bar{p}_{1,1}\bar{p}_{q,q'-1}\bar{p}_{u-1,u'}$$
$$- u\bar{p}_{q+1,q'}\bar{p}_{u-1,u'} - u'\bar{p}_{q,q'+1}\bar{p}_{u,u'-1}$$
$$- q\bar{p}_{u+1,u'}\bar{p}_{q-1,q'} - q'\bar{p}_{u,u'+1}\bar{p}_{q,q'-1} \quad \text{......................(xiv)}.$$

(xii) and (xiv) contain all the requisite data for the probable errors of random sampling in the case of two variables. They, of course, contain implicitly the case of one variable, as we have only to put q' and u' zero in order to fall back upon the formulae (vii) and (viii) of the first part of this paper (*Biometrika*, Vol. II. pp. 276—7).

(3) We may illustrate as follows:

(α) Correlation of errors in means:

$$r_{\bar{x}\bar{y}} = r_{xy} \quad \text{....................................(xv)}.$$

(β) Correlation of errors in standard deviations*:

$$N\sigma_{\bar{p}_{20}}\sigma_{\bar{p}_{02}}r_{\bar{p}_{20}\bar{p}_{02}} = \bar{p}_{22} - \bar{p}_{20}\bar{p}_{02},$$

$$\therefore \quad 4\sigma_x\sigma_y\sigma_{\sigma_x}\sigma_{\sigma_y}r_{\sigma_x\sigma_y} = \frac{\bar{p}_{22} - \bar{p}_{20}\bar{p}_{02}}{N},$$

$$\therefore \quad r_{\sigma_x\sigma_y} = \frac{\bar{p}_{22} - \bar{p}_{20}\bar{p}_{02}}{\sqrt{\bar{p}_{40} - \bar{p}^2_{20}}\sqrt{\bar{p}_{04} - \bar{p}^2_{02}}} \quad \text{........................(xvi)}.$$

This is the general value of the correlation between the standard deviations of two correlated variables. We may write it in the form

$$r_{\sigma_x\sigma_y} = \frac{\bar{p}_{22}/(\bar{p}_{20}\bar{p}_{02}) - 1}{\sqrt{(\beta_2 - 1)(\beta_2' - 1)}}$$

where β_2 and β_2' are the second β's for the two variables respectively

We may now investigate $\bar{p}_{22}$,

$$\bar{p}_{22} = \frac{1}{N}S\{(x_s - \bar{x})^2(y_{s'} - \bar{y})^2 n_{ss'}\}.$$

* We may safely write p_{22} for $p_{2,2}$ etc. when we leave the general formulae, but $p_{qq'-1}$ for $p_{q,q'-1}$ is capable of misinterpretation.

Let us sum for s' and keep x_s constant; then

$$\Sigma\{(y_{s'}-\bar{y})^2 n_{ss'}\} = \{n_s(\sigma^2_{y_s}+(\bar{y}_s-\bar{y})^2)\},$$

where σ_{y_s} is the standard deviation and $\bar{y}_s$ the mean of the array of y's for a given x_s. If the regression be linear and homoscedastic, then

$$\bar{y}_s-\bar{y}=r_{xy}\frac{\sigma_y}{\sigma_x}(x_s-\bar{x}),$$

and

$$\sigma^2_{y_s}=\sigma_y^2(1-r^2_{xy}).$$

Hence

$$\bar{p}_{22}=\frac{\sigma_r^2}{N}S\left\{n_s(x_s-\bar{x})^2(1-r^2_{xy})+n_s(x_s-\bar{x})^4\frac{r^2_{xy}}{\sigma_x^2}\right\}$$

$$=\sigma_x^2\sigma_y^2\{1-r^2_{xy}+r^2_{xy}\beta_2\},$$

$$\frac{\bar{p}_{22}}{\bar{p}_{20}\bar{p}_{02}}-1=r^2_{xy}(\beta_2-1).$$

Similarly

$$\frac{\bar{p}_{22}}{\bar{p}_{20}\bar{p}_{02}}-1=r^2_{xy}(\beta_2'-1),$$

on the assumption that the other regression is also linear and homoscedastic.

It is accordingly impossible for two variables to have a regression linear and homoscedastic in both senses unless β_2 has the same value for both variables. Clearly for most practical purposes we may take

$$\frac{\bar{p}_{22}}{\bar{p}_{20}\bar{p}_{02}}-1=r^2_{xy}\sqrt{(\beta_2-1)(\beta_2'-1)}.$$

Thus approximately

$$r_{\sigma_x\sigma_y}=r^2_{xy} \quad \text{.................................(xvii).}$$

This is identical with the result found by Pearson and Filon (*Phil. Trans.* Vol. 191, A, p. 242) on the assumption of normal correlation. It is now seen to be true, whenever we may assert linearity of regression and homoscedastic distribution for both variables.

(γ) Probable error of a coefficient of correlation.

Put

$$r_{xy}=\frac{\bar{p}_{11}}{\sqrt{\bar{p}_{20}\bar{p}_{02}}},$$

$$\frac{\delta r_{xy}}{r_{xy}}=\frac{\delta\bar{p}_{11}}{\bar{p}_{11}}-\frac{1}{2}\frac{\delta\bar{p}_{20}}{\bar{p}_{20}}-\frac{1}{2}\frac{\delta\bar{p}_{02}}{\bar{p}_{02}}.$$

Square, add for all random samples and divide by their number:

$$\frac{\sigma^2_{r_{xy}}}{r^2_{xy}}=\frac{1}{N}\left\{\frac{\bar{p}_{22}-\bar{p}_{11}^2}{\bar{p}_{11}^2}+\frac{1}{4}\frac{\bar{p}_{40}-\bar{p}_{20}^2}{\bar{p}_{20}^2}+\frac{1}{4}\frac{\bar{p}_{04}-\bar{p}_{02}^2}{\bar{p}_{02}^2}+\frac{1}{2}\frac{\bar{p}_{22}-\bar{p}_{20}\bar{p}_{02}}{\bar{p}_{20}\bar{p}_{02}}\right.$$
$$\left.-\frac{\bar{p}_{31}-\bar{p}_{11}\bar{p}_{20}}{\bar{p}_{11}\bar{p}_{20}}-\frac{\bar{p}_{13}-\bar{p}_{11}\bar{p}_{02}}{\bar{p}_{11}\bar{p}_{02}}\right\} \quad \text{...(xviii).}$$

This is the most general value of the standard deviation $\sigma_{r_{xy}}$ of a correlation coefficient. It was first given by Sheppard (*Phil. Trans.* Vol. 192, A, p. 128, with

an obvious printer's slip, the omission of r^2_{xy}). It is clear that we require a knowledge of at least an approximate value of p_{31} and p_{13} as well as p_{22} in order to simplify this expression, which is far too cumbersome for practical use.

Now
$$N \times \bar{p}_{31} = S\{n_{ss'}(x_s - \bar{x})^3 (y_{s'} - \bar{y})\}$$
$$= S\{n_s(\bar{y}_s - \bar{y})(x_s - \bar{x})^3\},$$
if we sum for the sth array of y's.

But, if the regression be linear,
$$\bar{y}_s - \bar{y} = \frac{r_{xy}\sigma_y}{\sigma_x}(x_s - \bar{x}),$$
therefore
$$N \times \bar{p}_{31} = \frac{r_{xy}\sigma_y}{\sigma_x} S\{n_s(x_s - \bar{x})^4\},$$
$$\bar{p}_{31} = r_{xy}\sigma_y\sigma_x{}^3\beta_2$$
$$= \bar{p}_{11}\bar{p}_{20}\beta_2 \quad\text{......(xix).}$$
Similarly
$$\bar{p}_{13} = r_{xy}\sigma_x\sigma_y{}^3\beta_2'$$
$$= \bar{p}_{11}\bar{p}_{02}\beta_2' \quad\text{......(xx).}$$

We can now substitute in (xviii), if we determine what value to give to p_{22}. If we take
$$\bar{p}_{22} = \sigma_x{}^2\sigma_y{}^2\{1 - r^2_{xy} + r^2_{xy} \times \tfrac{1}{2}(\beta_2 + \beta_2')\},$$
we have
$$\frac{\bar{p}_{22}}{\bar{p}_{20}\bar{p}_{02}} - 1 = r^2_{xy} \times \frac{\beta_2 - 1 + \beta_2' - 1}{2},$$
and
$$\frac{\bar{p}_{22}}{\bar{p}_{11}{}^2} = \frac{1}{r^2_{xy}} - 1 + \tfrac{1}{2}(\beta_2 + \beta_2'),$$
or
$$\frac{\bar{p}_{22}}{\bar{p}_{11}{}^2} - 1 = \frac{1 - r^2_{xy}}{r^2_{xy}} + \tfrac{1}{2}(\beta_2 - 1 + \beta_2' - 1).$$

Hence
$$\sigma^2_{r_{xy}} = \frac{r^2_{xy}}{N}\left\{\frac{1 - r^2_{xy}}{r^2_{xy}} + \tfrac{1}{2}(\beta_2 - 1 + \beta_2' - 1)(1 + \tfrac{1}{2}r^2_{xy}) \right.$$
$$\left. + \tfrac{1}{4}(\beta_2 - 1) + \tfrac{1}{4}(\beta_2' - 1) - (\beta_2 - 1) - (\beta_2' - 1)\right\}$$
$$= \frac{r^2_{xy}}{N}\left\{\frac{1 - r^2_{xy}}{r^2_{xy}} - \tfrac{1}{4}(\beta_2 - 1 + \beta_2' - 1)(1 - r^2_{xy})\right\}$$
$$= \frac{1 - r^2_{xy}}{N}\{1 - \tfrac{1}{4}(\beta_2 - 1 + \beta_2' - 1)r^2_{xy}\}$$
$$= \frac{1 - r^2_{xy}}{N}\{1 - r^2_{xy} - \tfrac{1}{4}(\beta_2 - 3 + \beta_2' - 3)r^2_{xy}\}.$$

Or
$$\sigma_{r_{xy}} = \frac{1 - r^2_{xy}}{\sqrt{N}}\left\{1 - \tfrac{1}{4}(\beta_2 - 3 + \beta_2' - 3)\frac{r^2_{xy}}{1 - r^2_{xy}}\right\}^{\frac{1}{2}} \quad\text{......(xxi).}$$

This result is of much interest. If the kurtosis be zero, then

$$\beta_2 - 3 = 0 = \beta_2' - 3,$$

and we have

$$\sigma_{r_{xy}} = \frac{1 - r^2_{xy}}{\sqrt{N}} \quad \text{.............................(xxii)},$$

the value originally given by Pearson and Filon for the standard deviation of a correlation coefficient when the frequency surface is Gaussian (*Phil. Trans.* Vol. 191, A, p. 242). We see accordingly that:

(i) Equal kurtosis is needful in the two variates if the regression is to be linear and the arrays to be homoscedastic in the case of each variable.

(ii) The ordinary value subject to (i) is only correct provided the kurtosis is zero, and this is true whether the distribution be Gaussian or not.

(iii) The ordinary formula may give very inaccurate results, if the kurtosis be considerable and the correlation high.

(iv) It is probable that (xxi), as we have taken a mean value for p_{22}, gives fairly good results even when the correlation is not linear.

Clearly we must always have

$$r_{xy} < \frac{2}{\sqrt{\beta_2 - 1 + \beta_2' - 1}} \quad \text{........................(xxiii)}.$$

Or, for linear regression in homoscedastic systems there is a superior limit to the correlation possible with given values of the kurtosis. This is an interesting point, and forms a remarkable limitation on the nature of double linear homoscedastic regression.

(4) We may now find the correlation between a product-moment $p_{q,q'}$ about fixed axes and $\bar{p}_{u,u'}$ a product-moment about axes through the centroid. We have to multiply together (vii) *bis* and (xiii). We have

$$\begin{aligned} N^2 \sigma_{p_{q,q'}} \sigma_{\bar{p}_{u,u'}} \times r_{p_{q,q'} \bar{p}_{u,u'}} = {} & S\left\{n_{ss'}\left(1 - \frac{n_{ss'}}{N}\right) x_s{}^q x_{s'}{}^{q'} \times (x_s - \bar{x})^u (y_{s'} - \bar{y})^{u'}\right\} \\ & - S\left\{\frac{n_{tt'} n_{ss'}}{N} x_t{}^q x_{t'}{}^{q'} \times (x_s - \bar{x})^u (y_{s'} - \bar{y})^{u'}\right\} \\ & - u\bar{p}_{u-1,u'} S\left\{\frac{n_{ss'}}{N}(x_s - \bar{x}) x_s{}^q y_{s'}{}^{q'}\right\} \\ & - u'\bar{p}_{u,u'-1} S\left\{\frac{n_{ss'}}{N}(y_{s'} - \bar{y}) x_s{}^q y_{s'}{}^{q'}\right\} \quad \text{.........(xxiv)}. \end{aligned}$$

We can simplify the form of (xxiv) by taking the fixed axes now through the centroid itself. This gives us

$$\sigma_{p_{q,q'}} \sigma_{\bar{p}_{u,u'}} r_{p_{q,q'} \bar{p}_{u,u'}} = \frac{1}{N}\{\bar{p}_{q+u,q'+u'} - \bar{p}_{q,q'}\bar{p}_{u,u'} - u\bar{p}_{u-1,u'}\bar{p}_{q+1,q'} - u'\bar{p}_{u,u'-1}\bar{p}_{q,q'+1}\}$$

$$\text{......(xxv)}.$$

Illustrations. (α) To find the correlation between a deviation due to random sampling in a mean and one in the standard deviation of the same variate.

Take $\quad q=1,\ q'=0, \quad \therefore\ p_{q,q'}=\bar{x}; \quad u=2,\ u'=0, \quad \therefore\ \bar{p}_{u,u'}=\mu_2.$

$$\sigma_{\bar{x}}\sigma_{\mu_2}r_{\bar{x}\mu_2}=\frac{1}{N}(\bar{p}_{30}-\bar{p}_{10}\bar{p}_{20}-2\bar{p}_{10}\bar{p}_{20}).$$

But $$\bar{p}_{10}=0, \quad \sigma_{\bar{x}}=\sigma_x/\sqrt{N},$$

$$\sigma_{\mu_2}=\sqrt{\mu_4-\mu_2^2}/\sqrt{N}=\mu_2\sqrt{\beta_2-1}/\sqrt{N}, \quad \bar{p}_{30}=\mu_3=\sqrt{\beta_1}\sigma_x^3.$$

Hence $$r_{\bar{x}\sigma_x}=r_{\bar{x}\mu_2}=\sqrt{\beta_1}/\sqrt{\beta_2-1} \quad \text{......................(xxvi)}.$$

This is perfectly general; we see that variations in the mean are independent of variations in the variability for all symmetrical systems including the Gaussian.

(β) To find the correlation between a deviation due to random sampling in the mean of one variate and one in the standard deviation of a correlated variate.

Take $\quad q=1,\ q'=0, \quad \therefore\ p_{q,q'}=\bar{x}; \quad u=0,\ u'=2, \quad \therefore\ \bar{p}_{u,u'}=\bar{p}_{02}=\mu_2'=\sigma_y^2.$

$$\sigma_{\bar{x}}\sigma_{\mu_2'}r_{\bar{x}\mu_2'}=\frac{1}{N}(\bar{p}_{12}-\bar{p}_{10}\bar{p}_{02}-2\bar{p}_{01}\bar{p}_{11}).$$

But $$\bar{p}_{10}=\bar{p}_{01}=0, \quad \sigma_{\bar{x}}=\sigma_x/\sqrt{N}, \quad \sigma_{\mu_2'}=\mu_2'\sqrt{\beta_2'-1}/\sqrt{N}.$$

It remains to consider

$$\bar{p}_{12}=\frac{1}{N}S\{n_{ss'}(x_s-\bar{x})(y_{s'}-\bar{y})^2\}$$

$$=\frac{1}{N}S\{n_s(\bar{x}_{s'}-\bar{x})(y_{s'}-\bar{y})^2\},$$

where $\bar{x}_{s'}$ = mean of array of x's corresponding to the $y_{s'}$ of y. Hence if the regression be linear

$$\bar{p}_{12}=\frac{1}{N}\frac{r_{xy}\sigma_x}{\sigma_y}S\{n_s(y_{s'}-\bar{y})^3\}$$

$$=\frac{r_{xy}\sigma_x}{\sigma_y}\mu_3'=r_{xy}\sigma_x\sigma_y^2\sqrt{\beta_1'}.$$

Thus we have $$r_{\bar{x}\sigma_y}=r_{\bar{x}\mu_2'}=r_{xy}\sqrt{\beta_1'}/\sqrt{\beta_2'-1} \quad \text{....................(xxvii)}.$$

Similarly $$r_{\bar{x}\sigma_x}=r_{xy}\sqrt{\beta_1}/\sqrt{\beta_2-1} \quad \text{......................(xxviii)}.$$

Clearly $$r_{\bar{x}\sigma_y}=r_{xy}\cdot r_{\bar{y}\sigma_y}, \quad =r_{\bar{x}\bar{y}}\cdot r_{\bar{y}\sigma_y},$$

and, $$r_{\bar{y}\sigma_x}=r_{xy}\cdot r_{\bar{x}\sigma_x}, \quad =r_{\bar{x}\bar{y}}\cdot r_{\bar{x}\sigma_x},$$

by (xv), which show us that these correlations are second order correlations, and proves that the correlation of the mean of one variate with the variability of a second is zero, for constant mean value of the second, since

$$_{\bar{y}}r_{\bar{x}\sigma_y}=\frac{r_{\bar{x}\sigma_y}-r_{\bar{x}\bar{y}}r_{\bar{y}\sigma_y}}{\sqrt{1-r^2_{\bar{x}\bar{y}}}\sqrt{1-r^2_{\bar{y}\sigma_y}}}.$$

(γ) To find the correlation between a mean and a coefficient of correlation, i.e. between $\bar{x}$ and r_{xy}.

We have to multiply $\delta\bar{x} = \delta p_{10}$ with

$$\frac{\delta r_{xy}}{r_{xy}} = \frac{\delta \bar{p}_{11}}{\bar{p}_{11}} - \frac{1}{2}\frac{\delta \bar{p}_{20}}{\bar{p}_{20}} - \frac{1}{2}\frac{\delta \bar{p}_{02}}{\bar{p}_{02}},$$

and apply (xxv) to each term. We have

$$\frac{1}{r_{xy}}\sigma_{r_{xy}}\sigma_{\bar{x}}r_{r_{xy}\bar{x}} = \frac{1}{N}\left\{\frac{\bar{p}_{21}}{\bar{p}_{11}} - \frac{1}{2}\frac{\bar{p}_{30}}{\bar{p}_{20}} - \frac{1}{2}\frac{\bar{p}_{12}}{\bar{p}_{02}}\right\}$$

$$= \frac{1}{N}\left\{\frac{r_{xy}\sigma_y\sigma_x^2\sqrt{\beta_1}}{r_{xy}\sigma_y\sigma_x} - \frac{1}{2}\frac{\sqrt{\beta_1}\sigma_x^3}{\sigma_x^2} - \frac{1}{2}\frac{r_{xy}\sigma_x\sigma_y^2\sqrt{\beta_1'}}{\sigma_y^2}\right\}$$

$$= \frac{1}{N}\sigma_x\left\{\tfrac{1}{2}(\sqrt{\beta_1} - r_{xy}\sqrt{\beta_1'})\right\},$$

using the values of $\bar{p}_{21}$ and $\bar{p}_{12}$ for linear regression. If we now use the value in (xxi) for $\sigma_{r_{xy}}$ we have

$$r_{r_{xy}\bar{x}} = \frac{\frac{1}{2}r_{xy}\{\sqrt{\beta_1} - r_{xy}\sqrt{\beta_1'}\}}{(1-r^2_{xy})\left\{1 - \frac{1}{4}(\beta_2 - 3 + \beta_2' - 3)\frac{r^2_{xy}}{1-r^2_{xy}}\right\}^{\frac{1}{2}}} \quad \text{.........(xxix)},$$

reducing when the kurtosis of both distributions is zero to

$$r_{r_{xy}\bar{x}} = \frac{1}{2}\frac{r_{xy}\{\sqrt{\beta_1} - r_{xy}\sqrt{\beta_1'}\}}{1-r^2_{xy}} \quad \text{.........................(xxx)},$$

and vanishing for all symmetrical linearly correlated variates, including of course Gaussian systems.

(δ) To find the correlation between a deviation in a standard deviation and one in a coefficient of correlation. We have to multiply $\delta\bar{p}_{20}$ by

$$\frac{\delta r_{xy}}{r_{xy}} = \frac{\delta \bar{p}_{11}}{\bar{p}_{11}} - \frac{1}{2}\frac{\delta \bar{p}_{20}}{\bar{p}_{20}} - \frac{1}{2}\frac{\delta \bar{p}_{02}}{\bar{p}_{02}}.$$

We find

$$\frac{1}{r_{xy}}\sigma_{\mu_2}\sigma_{r_{xy}}r_{\mu_2 r_{xy}} = \frac{\bar{p}_{31} - \bar{p}_{20}\bar{p}_{11}}{N\bar{p}_{11}} - \frac{1}{2}\frac{\bar{p}_{40} - \bar{p}_{20}^2}{N\bar{p}_{20}} - \frac{1}{2}\frac{\bar{p}_{22} - \bar{p}_{20}\bar{p}_{02}}{N\bar{p}_{02}}.$$

Hence, assuming linearity of regression, we may put

$$\bar{p}_{31}/\bar{p}_{11} = \sigma_x^2\beta_2, \quad \bar{p}_{40}/\bar{p}_{20} = \beta_2\sigma_x^2,$$

and approximately by the result immediately above Equation (xvii)

$$\bar{p}_{22}/\bar{p}_{02} - \bar{p}_{20} = \sigma_x^2 r^2_{xy}\sqrt{(\beta_2 - 1)(\beta_2' - 1)}.$$

Thus

$$\frac{1}{r_{xy}}\frac{\sigma_x^2\sqrt{\beta_2 - 1}}{\sqrt{N}}\sigma_{r_{xy}}r_{\mu_2 r_{xy}} = \frac{1}{N}\{\sigma_x^2(\beta_2 - 1) - \tfrac{1}{2}\sigma_x^2(\beta_2 - 1) - \tfrac{1}{2}\sigma_x^2 r^2_{xy}\sqrt{(\beta_2 - 1)(\beta_2' - 1)}\}.$$

Or
$$\sigma_{r_{xy}}r_{\sigma_x r_{xy}} = \frac{r_{xy}}{2\sqrt{N}}\{\sqrt{\beta_2 - 1} - r^2_{xy}\sqrt{\beta_2' - 1}\},$$

and by (xxi)

$$r_{\sigma_x r_{xy}} = \tfrac{1}{2} r_{xy} \frac{\sqrt{\beta_2 - 1} - r^2_{xy}\sqrt{\beta_2' - 1}}{(1 - r^2_{xy})\left\{1 - \tfrac{1}{4}(\beta_2 - 3 + \beta_2' - 3)\dfrac{r^2_{xy}}{1 - r^2_{xy}}\right\}^{\frac{1}{2}}} \quad \text{......(xxxi).}$$

Since homoscedastic linear regression supposes $\beta_2 = \beta_2'$ this result must be very close to

$$r_{\sigma_x r_{xy}} = \tfrac{1}{2} r_{xy}\sqrt{\beta_2 - 1}\bigg/\left(1 - \tfrac{1}{4}(\beta_2 - 3 + \beta_2' - 3)\frac{r^2_{xy}}{1 - r^2_{xy}}\right)^{\frac{1}{2}} \quad \text{......(xxxii)}$$

for such distributions.

For distributions in which the kurtosis is zero ($\beta_2 = 3$) we reach

$$r_{\sigma_x r_{xy}} = r_{xy}/\sqrt{2} \quad \text{..........................(xxxiii),}$$

a result already reached for Gaussian distributions by Pearson and Filon in 1897 *. It is now shown to be true for all homoscedastic, mesokurtic systems with linear regression.

(5) Two further probable errors are of interest. If we write $\bar{y}_x = ax + b$ for the regression line, what are the probable errors of

$$a = r_{xy}\sigma_y/\sigma_x \text{ and } b = \bar{y} - r_{xy}\sigma_y\bar{x}/\sigma_x\,?$$

It will be sufficient to give the values of σ_a and σ_b when the frequencies are symmetrical and the regression linear. In this case

$$r_{\bar{x}\sigma_x} = r_{\bar{y}\sigma_y} = r_{\bar{x}\sigma_y} = r_{\bar{y}\sigma_x} = 0, \quad r_{\bar{x}\bar{y}} = r_{xy},$$

$$r_{r_{xy}\bar{x}} = r_{r_{xy}\bar{y}} = 0, \quad r_{r_{xy}\sigma_x} = r_{r_{xy}\sigma_y} = \frac{1}{\sqrt{2}} r_{xy}, \quad r_{\sigma_x\sigma_y} = r^2_{xy}.$$

Writing $\tau_x = \bar{x}/\sigma_x$, we have

$$\frac{\delta a}{a} = \frac{\delta\sigma_y}{\sigma_y} + \frac{\delta r_{xy}}{r_{xy}} - \frac{\delta\sigma_x}{\sigma_x},$$

and $$\delta b = \delta\bar{y} - a\delta\bar{x} - \delta r_{xy}\sigma_y\tau_x - r_{xy}\tau_x\delta\sigma_y + a\tau_x\delta\sigma_x.$$

Whence proceeding in the usual manner we deduce

$$\sigma_a = \frac{\sigma_y}{\sigma_x}\sqrt{1 - r^2_{xy}}\Big/\sqrt{N} \quad \text{................................(xxxiv)†,}$$

$$\sigma_b = \sigma_y\sqrt{1 - r^2_{xy}}\sqrt{1 + \tau_x^2}/\sqrt{N} = \sigma_a\sqrt{\bar{x}^2 + \sigma_x^2} \quad \text{........(xxxv).}$$

These enable us to determine any significant difference between two regression lines.

We can go a stage further and ask what is the probable error of the mean of an array, $\bar{y}_x$, as found from the regression line. We have

$$\delta\bar{y}_x = x\delta a + \delta b.$$

* *Phil. Trans.* Vol. 191, A, p. 242.

† Cf. *Phil. Trans.* Vol. 191, A, p. 245.

We require accordingly $\sigma_a \sigma_b r_{ab}$. Multiply the above expressions for δa and δb together, summing and dividing by the number of samples we find

$$\sigma_a \sigma_b r_{ab} = -\frac{r_x \sigma_y^2}{N\sigma_x}(1 - r^2_{xy}) \quad \text{......................(xxxvi).}$$

Whence we deduce
$$\sigma_{\bar{y}_x} = \frac{\sigma_y \sqrt{1 - r^2}}{\sqrt{N}} \left\{1 + \frac{(x - \bar{x})^2}{\sigma_x^2}\right\}^{\frac{1}{2}} \quad \text{..................(xxxvii).}$$

We note the increase of inaccuracy of the means of arrays far from the mean of the whole population*.

The above equations embrace the chief results for the probable errors and errors' correlations of systems of two variables. They have been reached independently of any system of Gaussian distribution.

* Results (xxxv), (xxxvi) and (xxxvii) are published here for the first time.

THE RELATIONSHIP BETWEEN THE WEIGHT OF THE SEED PLANTED AND THE CHARACTERISTICS OF THE PLANT PRODUCED. I*.

By J. ARTHUR HARRIS, Ph.D., Carnegie Institution of Washington.

I. Introductory Remarks.

In practical agriculture the quality of the seed planted is universally recognized as of fundamental importance. Three requisites are essential: freedom from noxious impurities, purity of breed and viability. All three of these points have attracted close attention, and have a voluminous literature. But given seeds belonging to the required variety, free from undesirable impurities and germinating successfully, it seems of theoretical interest at least, and perhaps of much practical importance as well, to ascertain the degree of relationship between the size of the seed planted and the characteristics of the individual into which it develops.

This question although by no means so extensively discussed as the others has received considerable attention. Much has been written concerning the desirability of winnowing seed to remove the lighter grains. A review of the literature of this field would lead us too far from our immediate purpose, which is to present the data derived from some rather extensive quantitative investigations.

It will not, however, be unprofitable to call attention to certain general deficiencies of the previous work, especially since this will define the point of view directing the studies described here.

(1) In many cases the distinction between perfectly matured but small seeds and potentially large but immature, blighted or shrivelled seeds has been disregarded.

(2) The method of grading the seed has, generally speaking, been neither uniform nor logical. Usually the separation has been only into "heavy" and

* This first paper is limited to the presentation of the data for number of pods per plant in twenty series of garden beans belonging to three different varieties. Much material for other varieties is in hand. The constants for other characters, e.g. number of ovules and number of seeds per pod and seed weight, are being calculated and will be presented later with more general discussions.

"light," or at most into "heavy," "medium" and "light." The meaning of such crudely defined terms of course differs from experiment to experiment. They may, with certain limitations, enable one to say which class of seeds gives the best results. They do not permit comparisons of the advantages to be gained by seed selection in different varieties. They do not allow of the writing of a general formula enabling one to predict the yield from seed of given weight. Yet such a formula is precisely what is needed. The advantage of sorting seed—if there be an advantage—depends not only upon the increased yield (or increased uniformity of the crop in certain cases) but also upon the cost of carrying out such a selection. In deciding upon the stringency which is profitable in the seed selection, the practical breeder should know the exact weight to be attached to each factor.

(3) The experiments have not been carried out in a way to make possible the calculation of the probable errors of the results. In many cases the experiments have been small, and the conclusions are open to serious question, both on the ground of possible experimental errors and on the ground of the probable errors of random sampling.

All of these difficulties can be overcome by the application of the modern statistical methods to the problems. A wide series of such biometric constants deduced from carefully conducted physiological investigations ought to be of great service to the man dealing with the practical problems of agriculture.

II. Material and Methods.

The materials for this study are drawn from three varieties of garden beans—the White Flageolet, White Navy and Ne Plus Ultra. Altogether there are twenty series grown under most diverse environmental conditions. Their history has already been given* for another purpose and need not be repeated here. The symbols used to designate the different lots are the same in both papers; hence the reader may make any comparisons which he sees fit.

All of the seeds planted were, as far as could be determined by inspection, perfect in form and development. Each was individually weighed and classified in a uniformly graduated scale. These individually weighed and individually labelled seeds were then mixed and planted at random in rows by varieties, the rows in their turn being scattered over the field, more or less at random, in order to counteract by chance distribution any influence of the possible heterogeneity of the substratum upon the characters of the plants. All the determinations of number of pods produced were made on individual plants.

Thus the probable errors for the statistical constants for weight of seed planted, number of pods per plant, and for the correlation between the weight of

* Harris, J. Arthur: "A First Study of the Influence of the Starvation of the Ascendants upon the Characteristics of the Descendants." *Amer. Nat.*, 1912. In press.

the seed planted and number of pods per plant, can be determined. The correlations between weight of seed planted and number of pods produced are comparable from variety to variety or from cultural condition to cultural condition.

To render the results more intelligible, the straight line regression equations have also been calculated and represented in a series of diagrams. These show by the slope of the lines the (smoothed) change in mean number of pods per plant associated with changes in the weight of the seed planted.

The reader unacquainted with higher statistical methods need only remember that the coefficient of correlation describes the degree of interdependence between two variables on a scale of -1 to $+1$. This measure of interdependence is, therefore, quite independent of the magnitude and of the variability of either or both of the characters in question. The regression coefficient, on the other hand, shows the absolute amount of change in a second character y consequent upon a change of one unit on the scale of the first character x. Concretely, in our present case, the regression coefficient shows the absolute increase (or decrease) in number of pods per plant associated with an increase (or decrease) of one unit in the weight of the seed planted. "Increase" or "decrease" is measured from the average condition in the population as a whole.

The correlation coefficient is fully justified as a measure of interdependence only when regression is linear, that is to say, when the mean value of y increases at a uniform rate throughout the whole range of x. Where regression is not strictly linear, the coefficient of correlation still furnishes in many cases a very satisfactory measure of the intensity of relationship between two variables. This is true in the present case.

All the weighings were made on seeds which had dried for several months at laboratory temperature. Drying at high temperatures was of course precluded by the fact that the seeds were to be used for planting. Drying in a vacuum over sulphuric acid could not be undertaken because of the excessive labour involved where each seed had to be followed individually throughout the whole work. The weight unit adopted was ·025 gram. Hence to obtain means and standard deviations of weights in grams deduct ·5 from values in tables and multiply by ·025.

The correlation tables showing the relationship between the weight of seed planted and the number of pods produced are entirely too bulky for publication. It is possible, however, to present the essential data by showing the total number of pods produced by each grade of seed weight (Tables III—VI). A convenient method of calculation for such cases has been suggested elsewhere*.

In deducing the correlations from such tables, the means and standard deviations for the two characters involved are required. The distributions of numbers of pods per plant for the twenty series have already been published† for a quite

* Harris, J. Arthur: "The Arithmetic of the Product Moment Method of Calculating the Coefficient of Correlation." *Amer. Nat.* Vol. XLIV. pp. 693—699, 1910.

† *Amer. Nat.* 1912. In press.

different purpose. The tables are very large and need not be given here. The physical constants deduced from them appear in Table I. The distribution of the seed weight is shown in the frequency columns of the reduced correlation tables. The physical constants* for seed weight are given in Table II.

III. Analysis of Data.

1. Number of pods per plant in Navy, White Flageolet and Ne Plus Ultra†.

From Tables III to VI the value of the rough product moment $\Sigma(w'p')$ about 0 as origin may be calculated straight away by multiplying up the total pods by the number of the weight class (in parentheses) and summing. The coefficient of correlation, r, is then deduced from the formula

$$r_{wp} = \frac{\Sigma(w'p')/N - \bar{w}\bar{p}}{\sigma_w \sigma_p},$$

while the equation for the regression straight line is given by

$$p = \left(\bar{p} - r_{wp}\frac{\sigma_p}{\sigma_w}\bar{w}\right) - r_{wp}\frac{\sigma_p}{\sigma_w}w,$$

where p and w represent weight of seeds and pods per plant, the bars denote the population means of the respective characters, and the sigmas their standard deviations. The variable p is integral and there is no need for grouping; the unit of w is ·025 gram, with class 1 ranging from 0 to ·025, and centered at ·0125 gram.

The correlation coefficients and regression equations are given in Table VII.

The straight line regression equations being available for all series, it is only necessary to determine the empirical means for number of pods per plant to test by graphical methods the linearity of the regression of pods per plant on weight of seed planted.

It is not feasible—because of the rather great labour demanded, and the complexity of the diagrams—to do this for all the series. In the diagrams (Figures 1—3) the slope of all the twenty lines given by the equations is shown. In the case of certain of these lines, selected quite at random, the values of the empirical means are indicated in the usual way.

The empirical means are scattered with some irregularity and from inspection alone one might suspect that regression is not strictly linear‡. In other words,

* Sheppard's correction was applied in calculating the moments for seed weight, but not in obtaining those for number of pods per plant, since the latter varies discretely.

† Other series of material, and the relationships for seed weight and characters other than number of pods are being reduced.

‡ The sensitiveness of the number of pods per plant to environmental conditions (and consequently the great variability in the means, especially near the ends of the range of seed weight) is so great that it has seemed inadvisable to attempt more refined mathematical treatment of the problem of linearity of regression. I hope to do this later on larger series of material.

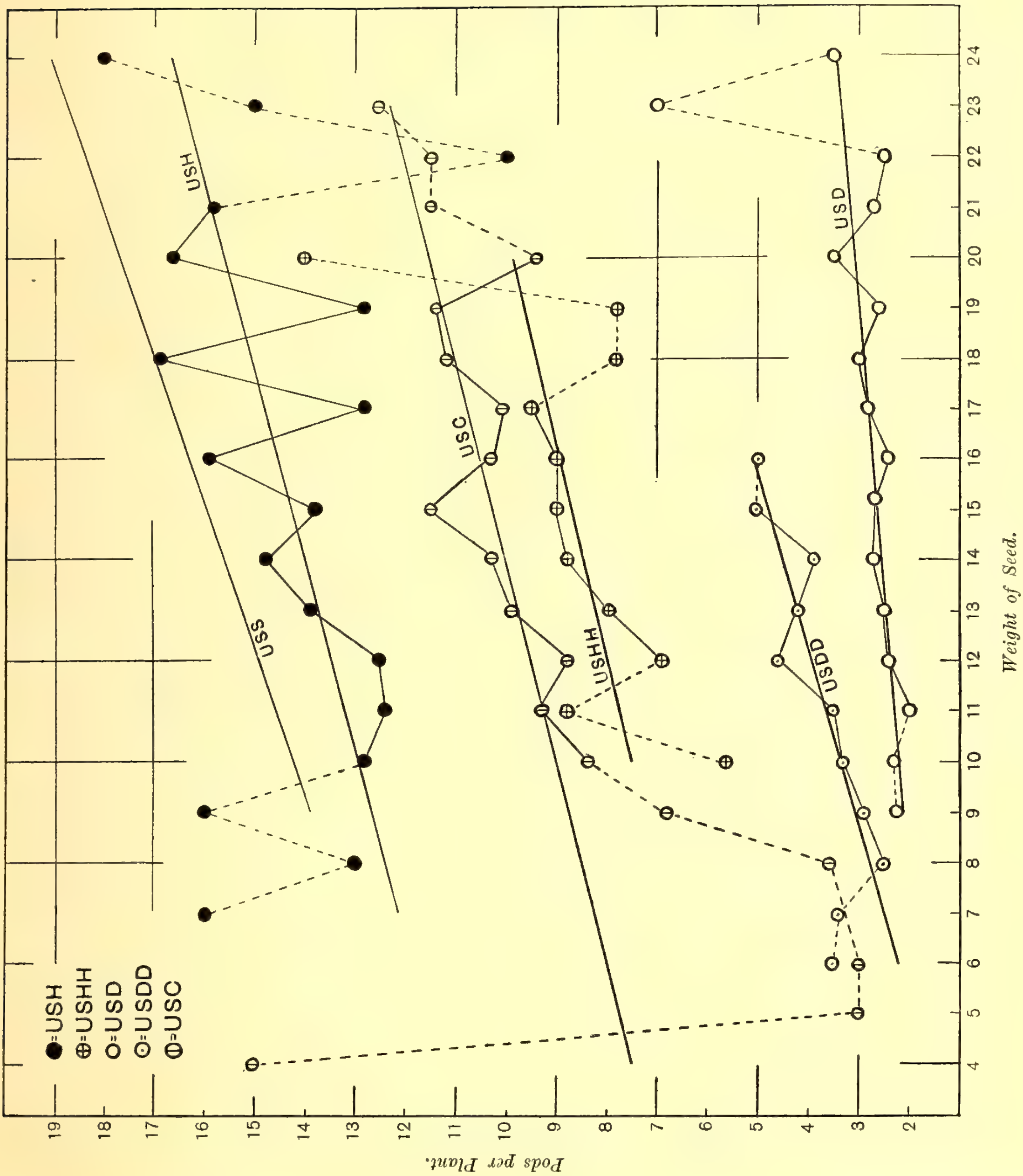

● = USH
⊕ = USHH
O = USD
⊙ = USDD
Φ = USC
USH
USS
USC
USHH
USD
USDD
Weight of Seed.
Pods per Plant.

Fig. 2.

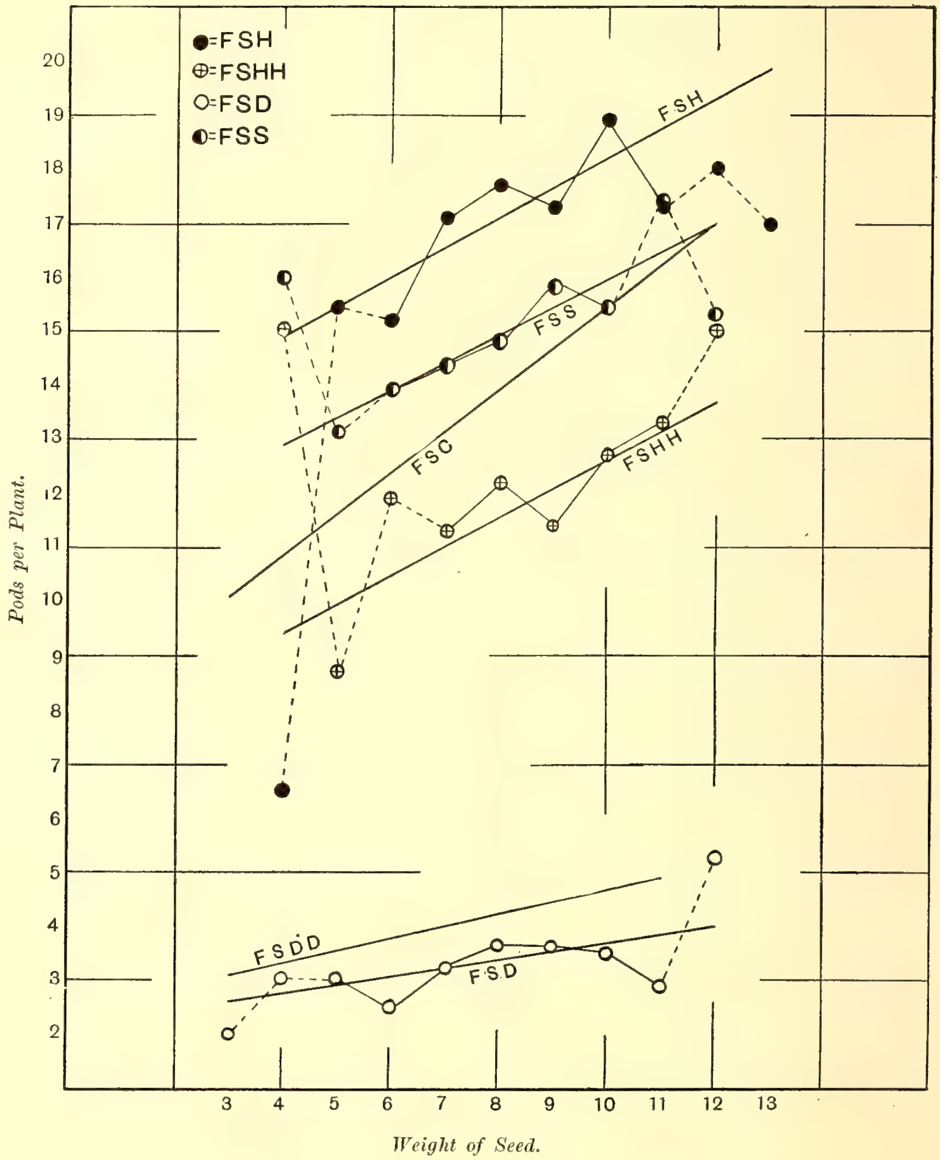
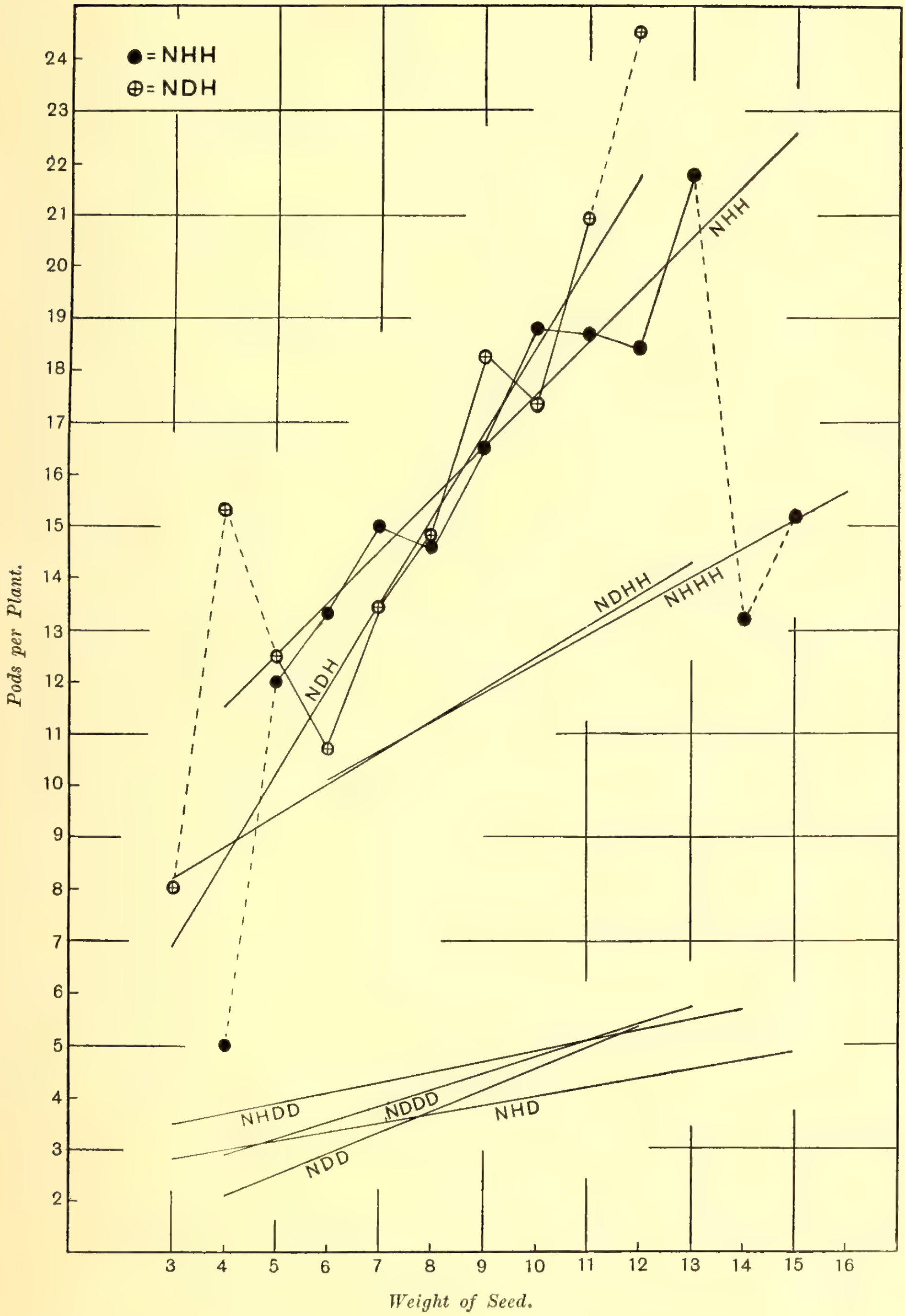

●=FSH
⊕=FSHH
○=FSD
◐=FSS
FSH
FSS
FSC
FSHH
FSDD
FSD
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
Pods per Plant.
3
4
5
6
7
8
9
10
11
12
13
Weight of Seed.

Fig. 3.


● = NHH
⊕ = NDH
NHH
NDH
NDHH
NHHH
NHDD
NDDD
NHD
NDD
Pods per Plant.
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Weight of Seed.

the mean number of pods may not increase at the same rate from the lowest to the highest, the lightest to the heaviest, grade of seeds. Practically speaking they may be considered to do so.

The coefficients of correlation are rather small, ranging from about ·120 to ·280. The average of the twenty series is ·1615. It may be noted, however, that they are in every instance positive, thus indicating that the selection of larger seeds will give a somewhat higher yield of pods.

The second term of the regression equation enables us to read off at once the increase in the number of pods per plant to be secured by selecting seeds one unit (i.e. 25 mg.) or more above the average. These values range widely, as is to be expected from the fact that the crops upon which they were calculated were purposely grown under most diverse cultural conditions. The slopes of the regression lines will make clear to the eye the advantages to be secured by planting heavier seed.

COLD SPRING HARBOR, U.S.A.,
January 29, 1912.

TABLE I.

Pods per Plant.

Series	Total Plants	Mean and Probable Error	Standard Deviation and Probable Error	Coefficient of Variation and Probable Error
NHH ...	1484	16·99 ± ·15	8·67 ± ·11	51·02 ± 0·78
NHHH ...	1271	11·93 ± ·10	5·17 ± ·07	43·29 ± 0·68
NHD ...	1416	3·97 ± ·03	1·94 ± ·02	48·97 ± 0·76
NHDD ...	1204	4·58 ± ·05	2·38 ± ·03	51·84 ± 0·88
NDD ...	513	3·59 ± ·06	1·89 ± ·04	52·66 ± 1·38
NDDD ...	459	4·41 ± ·06	1·93 ± ·04	43·86 ± 1·15
NDH ...	670	14·62 ± ·21	8·24 ± ·15	56·39 ± 1·33
NDHH ...	565	11·83 ± ·14	4·96 ± ·10	41·94 ± 0·98
FSS ...	868	15·03 ± ·17	7·41 ± ·12	49·34 ± 0·97
FSC ...	586	14·22 ± ·21	7·38 ± ·15	51·90 ± 1·27
FSH ...	475	17·29 ± ·25	7·94 ± ·17	45·89 ± 1·20
FSHH ...	429	11·84 ± ·16	4·80 ± ·11	40·50 ± 1·08
FSD ...	428	3·43 ± ·06	1·69 ± ·04	49·42 ± 1·39
FSDD ...	387	4·04 ± ·06	1·73 ± ·04	42·85 ± 1·22
USS ...	680	15·74 ± ·16	6·04 ± ·11	38·37 ± 0·80
USC ...	530	10·14 ± ·13	4·38 ± ·09	43·23 ± 1·05
USH ...	361	14·04 ± ·20	5·55 ± ·14	39·56 ± 1·14
USHH ...	224	8·44 ± ·15	3·24 ± ·10	38·45 ± 1·25
USD ...	312	2·59 ± ·05	1·23 ± ·03	47·30 ± 1·54
USDD ...	237	3·62 ± ·09	2·10 ± ·07	57·97 ± 2·32

TABLE II.

Seed Weights in Working Scale.

Series	Total Seeds	Mean and Probable Error	Standard Deviation and Probable Error	Coefficient of Variation and Probable Error
NHH ...	1484	9·44 ± ·03	1·51 ± ·02	15·96 ± ·20
NHHH ...	1271	9·27 ± ·03	1·33 ± ·02	14·36 ± ·20
NHD ...	1416	9·46 ± ·03	1·51 ± ·02	15·93 ± ·21
NHDD ...	1204	8·46 ± ·03	1·41 ± ·02	16·66 ± ·24
NDD ...	513	7·73 ± ·04	1·32 ± ·03	17·05 ± ·37
NDDD ...	459	8·76 ± ·04	1·30 ± ·03	14·89 ± ·34
NDH ...	670	7·62 ± ·03	1·28 ± ·02	16·83 ± ·32
NDHH ...	565	8·93 ± ·04	1·24 ± ·03	13·93 ± ·28
FSS ...	868	8·20 ± ·03	1·43 ± ·02	17·48 ± ·29
FSC ...	586	8·31 ± ·04	1·40 ± ·03	16·81 ± ·34
FSH ...	475	8·33 ± ·04	1·44 ± ·03	17·28 ± ·39
FSHH ...	429	8·48 ± ·04	1·07 ± ·02	12·67 ± ·30
FSD ...	428	8·22 ± ·05	1·43 ± ·03	17·38 ± ·41
FSDD ...	387	7·19 ± ·04	1·04 ± ·03	14·52 ± ·36
USS ...	680	14·14 ± ·07	2·71 ± ·05	19·18 ± ·36
USC ...	530	14·50 ± ·08	2·58 ± ·05	17·82 ± ·38
USH ...	361	14·33 ± ·09	2·65 ± ·07	18·50 ± ·48
USHH ...	224	13·94 ± ·09	1·93 ± ·06	13·82 ± ·45
USD ...	312	14·39 ± ·10	2·63 ± ·07	18·28 ± ·51
USDD ...	237	10·85 ± ·08	1·78 ± ·06	16·37 ± ·52

TABLE III.

Weight of Seed Planted	Series *NHH*		Series *NHHH*		Series *NHD*		Series *NHDD*	
	F	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods
·050—·075 (3)	—	—	—	—	1	2	2	4
·075—·100 (4)	1	5	—	—	—	—	3	9
·100—·125 (5)	3	36	—	—	3	10	14	54
·125—·150 (6)	23	306	10	89	18	57	80	350
·150—·175 (7)	96	1439	95	1015	89	294	230	1017
·175—·200 (8)	281	4102	265	2932	276	1019	319	1382
·200—·225 (9)	401	6614	376	4583	386	1559	285	1317
·225—·250 (10)	330	6200	308	3701	298	1235	157	782
·250—·275 (11)	216	4040	151	1959	211	860	80	401
·275—·300 (12)	94	1734	51	668	95	413	25	153
·300—·325 (13)	25	545	8	106	28	117	6	29
·325—·350 (14)	9	119	4	63	8	40	3	19
·350—·375 (15)	5	76	2	39	3	13	—	—
·375—·400 (16)	—	—	1	9	—	—	—	—
Totals ...	1484	25216	1271	15164	1416	5619	1204	5517

TABLE IV.

Weight of Seed Planted	Series *NDD*		Series *NDDD*		Series *NDH*		Series *NDHH*	
	F	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods
·050—·075 (3)	—	—	—	—	1	8	1	2
·075—·100 (4)	3	6	2	4	8	122	—	—
·100—·125 (5)	25	65	4	17	26	326	1	17
·125—·150 (6)	53	152	15	47	76	814	13	116
·150—·175 (7)	143	481	44	195	198	2653	42	427
·175—·200 (8)	151	564	124	524	209	3103	160	1847
·200—·225 (9)	87	341	147	616	104	1890	164	1963
·225—·250 (10)	42	175	87	420	37	641	130	1613
·250—·275 (11)	8	55	27	140	9	188	42	546
·275—·300 (12)	1	4	6	42	2	49	9	123
·300—·325 (13)	—	—	3	18	—	—	3	28
·325—·350 (14)	—	—	—	—	—	—	—	—
·350—·375 (15)	—	—	—	—	—	—	—	—
·375—·400 (16)	—	—	—	—	—	—	—	—
Totals ...	513	1843	459	2023	670	9794	565	6682

TABLE V.

Weight of Seed Planted	Series *FSS*		Series *FSC*		Series *FSH*		Series *FSHH*		Series *FSD*		Series *FSDD*	
	F	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods
·050—·075 (3)	—	—	1	20	—	—	—	—	2	4	1	3
·075—·100 (4)	6	96	1	5	2	13	1	15	3	9	2	4
·100—·125 (5)	21	275	15	173	9	139	3	26	10	30	10	40
·125—·150 (6)	72	1002	40	463	43	652	13	155	25	63	79	286
·150—·175 (7)	170	2432	98	1238	80	1369	43	445	78	252	166	676
·175—·200 (8)	241	3566	168	2355	116	2049	166	2023	138	498	84	356
·200—·225 (9)	194	3065	154	2427	131	2265	131	1493	102	364	36	155
·225—·250 (10)	118	1827	72	1164	64	1207	60	760	43	152	8	41
·250—·275 (11)	36	627	31	420	24	414	10	133	20	57	1	1
·275—·300 (12)	10	153	6	69	5	90	2	30	7	37	—	—
·300—·325 (13)	—	—	—	—	1	17	—	—	—	—	—	—
Totals ...	868	13043	586	8334	475	8215	429	5080	428	1466	387	1562

TABLE VI.

Weight of Seed Planted	Series *USS*		Series *USC*		Series *USH*		Series *USHH*		Series *USD*		Series *USDD*	
	F	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods	*F*	Total Pods
·075—·100 (4)	—	—	1	15	—	—	—	—	—	—	—	—
·100—·125 (5)	—	—	1	3	—	—	—	—	—	—	—	—
·125—·150 (6)	—	—	1	3	—	—	—	—	—	—	2	7
·150—·175 (7)	—	—	—	—	1	16	—	—	—	—	5	17
·175—·200 (8)	—	—	2	7	1	13	—	—	—	—	13	33
·200—·225 (9)	8	98	4	27	1	16	—	—	5	11	31	91
·225—·250 (10)	18	236	5	42	10	128	5	28	3	7	49	161
·250—·275 (11)	57	943	27	252	27	336	20	176	16	32	61	216
·275—·300 (12)	118	1745	62	543	43	536	26	180	48	114	34	155
·300—·325 (13)	124	1924	92	916	73	1020	43	346	55	135	22	93
·325—·350 (14)	117	1764	109	1120	56	828	49	434	60	163	14	55
·350—·375 (15)	79	1261	70	805	61	845	36	323	53	142	5	25
·375—·400 (16)	43	731	60	616	28	445	26	235	19	47	1	5
·400—·425 (17)	34	535	28	284	15	192	8	76	14	39	—	—
·425—·450 (18)	19	295	21	236	12	203	6	47	9	28	—	—
·450—·475 (19)	23	402	19	216	11	141	4	31	12	31	—	—
·475—·500 (20)	16	323	16	151	8	133	1	14	8	28	—	—
·500—·525 (21)	13	246	8	92	11	174	—	—	3	8	—	—
·525—·550 (22)	6	94	2	23	1	10	—	—	4	10	—	—
·550—·575 (23)	3	57	2	25	1	15	—	—	1	7	—	—
·575—·600 (24)	2	48	—	—	1	18	—	—	2	7	—	—
Totals ...	680	10702	530	5376	361	5069	224	1890	312	809	237	858

TABLE VII.

Number of Pods and Weight of Seed in Working Scale of Weights.

Series	Coefficient of Correlation and Probable Error	Regression Straight Line Equation
NHH ...	·177 ± ·017	$p = 7{\cdot}383 + 1{\cdot}017\,w$
NHHH ...	·145 ± ·019	$p = 6{\cdot}719 + 0{\cdot}562\,w$
NHD ...	·129 ± ·018	$p = 2{\cdot}396 + 0{\cdot}166\,w$
NHDD ...	·121 ± ·019	$p = 2{\cdot}862 + 0{\cdot}203\,w$
NDD ...	·282 ± ·027	$p = {\cdot}469 + 0{\cdot}404\,w$
NDDD ...	·215 ± ·030	$p = 1{\cdot}618 + 0{\cdot}319\,w$
NDH ...	·258 ± ·024	$p = 1{\cdot}982 + 1{\cdot}657\,w$
NDHH ...	·152 ± ·028	$p = 6{\cdot}412 + 0{\cdot}606\,w$
FSS ...	·098 ± ·023	$p = 10{\cdot}856 + 0{\cdot}508\,w$
FSC ...	·147 ± ·027	$p = 7{\cdot}787 + 0{\cdot}774\,w$
FSH ...	·100 ± ·031	$p = 12{\cdot}692 + 0{\cdot}553\,w$
FSHH ...	·121 ± ·032	$p = 7{\cdot}256 + 0{\cdot}541\,w$
FSD ...	·130 ± ·032	$p = 2{\cdot}164 + 0{\cdot}153\,w$
FSDD ...	·144 ± ·034	$p = 2{\cdot}328 + 0{\cdot}238\,w$
USS ...	·155 ± ·025	$p = 10{\cdot}876 + 0{\cdot}344\,w$
USC ...	·150 ± ·029	$p = 6{\cdot}453 + 0{\cdot}255\,w$
USH ...	·129 ± ·035	$p = 10{\cdot}165 + 0{\cdot}271\,w$
USHH ...	·143 ± ·044	$p = 5{\cdot}076 + 0{\cdot}241\,w$
USD ...	·195 ± ·037	$p = 1{\cdot}287 + 0{\cdot}091\,w$
USDD ...	·241 ± ·041	$p = {\cdot}535 + 0{\cdot}284\,w$

ON THE PROBABLE ERROR OF A COEFFICIENT OF CORRELATION AS FOUND FROM A FOURFOLD TABLE.

By KARL PEARSON, F.R.S.

LET the fourfold table be

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	N

Then on the assumption that the frequency distribution is normal, we can by aid of Everitt's Tables of the Tetrachoric Functions* rapidly find r. I have shown in a paper published in the *Phil. Trans.* in 1900† that found in this way

Probable error of r

$$= \frac{\cdot 67449}{\sqrt{N}\chi_0}\left\{\frac{(a+d)(c+b)}{4N^2} + \psi_2^2\frac{(a+c)(d+b)}{N^2} + \psi_1^2\frac{(a+b)(d+c)}{N^2} + 2\psi_1\psi_2\frac{ad-bc}{N^2} - \psi_2\frac{ab-cd}{N^2} - \psi_1\frac{ac-bd}{N^2}\right\}^{\frac{1}{2}} \quad \ldots \text{(i)},$$

where
$$\psi_1 = \frac{1}{\sqrt{2\pi}}\int_0^{\beta_1} e^{-\frac{1}{2}z^2}dz, \qquad \psi_2 = \frac{1}{\sqrt{2\pi}}\int_0^{\beta_2} e^{-\frac{1}{2}z^2}dz,$$

$$\beta_1 = \frac{h-rk}{\sqrt{1-r^2}}, \qquad \beta_2 = \frac{k-rh}{\sqrt{1-r^2}},$$

$$\chi_0 = \frac{1}{2\pi}\frac{1}{\sqrt{1-r^2}}e^{-\frac{1}{2}\frac{1}{1-r^2}(h^2+k^2-2rhk)},$$

* *Biometrika*, Vol. VII, p. 436, and Vol. VIII, p. 385.

† *Phil. Trans.* Vol. 195 A, p. 14. Owing to the carelessness of the printers my χ_0 was put as $\sqrt{\chi_0}$ and the last N^2 in the denominator as N_2.

and h and k have their usual meaning defined by the integrals

$$\frac{(a+c)-(b+d)}{2N}=\frac{1}{\sqrt{2\pi}}\int_0^h e^{-\frac{1}{2}z^2}dz=\tfrac{1}{2}\alpha_1,\ \text{say};$$

$$\frac{(a+b)-(c+d)}{2N}=\frac{1}{\sqrt{2\pi}}\int_0^k e^{-\frac{1}{2}z^2}dz=\tfrac{1}{2}\alpha_2,\ \text{say}.$$

Let $H=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}h^2}$, $K=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}k^2}$ as usual.

Now the formula (i) above for the probable error of r is admittedly laborious in use. I have tried in many ways, while retaining its full accuracy, to throw it into a form involving less laborious calculations; I have not succeeded, however, in achieving any sensible reduction in its complexity, as long as I maintain its complete generality.

Although many hundred fourfold tables have now been published, many of which give such small correlations that their true significance can only be settled by a knowledge of their probable errors, yet I find only 40 to 50 probable errors have so far been determined. This matter seems so regrettable that I have sought for a fairly easy method of determining a closely empirical expression for the probable error of r which is likely to be of service, and can be adapted easily to tables.

I consider first two extreme cases. If h and k are both zero, or the fourfold division at the mean, then $\psi_1=\psi_2=0$*,

Probable error of r

$$=\frac{\cdot 67449\ 2\pi\sqrt{1-r^2}}{\sqrt{N}}\left\{\frac{(a+d)(b+c)}{4N^2}\right\}^{\frac{1}{2}}=\frac{\cdot 67449\ \sqrt{1-r^2}}{\sqrt{N}}\frac{\pi}{2}\left\{\frac{16ab}{N^2}\right\}^{\frac{1}{2}},$$

since in this case $a=d$, and $b=c$.

But for a division at the mean by Sheppard's Theorem

$$r=\cos\pi\frac{b}{a+b}=\sin\left(\frac{\pi}{2}-\frac{\pi b}{a+b}\right),$$

or

$$(\sin^{-1}r)/\tfrac{1}{2}\pi=(a-b)/(a+b).$$

Hence

$$1-\left(\frac{\sin^{-1}r}{\frac{1}{2}\pi}\right)^2=\frac{4ab}{(a+b)^2}=\frac{16ab}{N^2}.$$

Substituting we have:

$$\text{Probable error of } r=\frac{\cdot 67449}{\sqrt{N}}\frac{\pi}{2}\sqrt{1-r^2}\sqrt{1-\left(\frac{\sin^{-1}r}{90^\circ}\right)^2}\ \ldots\ldots\text{(ii)},$$

if the angle of the inverse sine be read in degrees.

Again if $r=0$, the probable error of r may be obtained from (i) whatever the values of h and k. For in this case

$$ad-bc=0,\quad \psi_1=\tfrac{1}{2}\alpha_1,\quad \psi_2=\tfrac{1}{2}\alpha_2,\quad \chi_0=HK.$$

* *Phil. Trans.* Vol. 192 A, p. 141 and Vol. 195 A, p. 7.

We have $(b+d)/N = \frac{1}{2}(1-\alpha_1), \quad (a+c)/N = \frac{1}{2}(1+\alpha_1),$

$$(a+b)/N = \tfrac{1}{2}(1+\alpha_2), \quad (c+d)/N = \tfrac{1}{2}(1-\alpha_2),$$

$$\frac{a+d}{N} = \frac{(a+b)(a+c)}{N^2} + \frac{ad-bc}{N^2} + \frac{(c+d)(b+d)}{N^2} + \frac{ad-bc}{N^2}$$

$$= \tfrac{1}{4}(1+\alpha_2)(1+\alpha_1) + \tfrac{1}{4}(1-\alpha_2)(1-\alpha_1) = \tfrac{1}{2}(1+\alpha_1\alpha_2),$$

since $ad - bc = 0$ in the original population.

Similarly: $$\frac{c+b}{N} = \tfrac{1}{2}(1-\alpha_1\alpha_2).$$

$$\frac{ab-cd}{N^2} = \frac{a(N-a-c-d)-cd}{N^2}$$

$$= \frac{a}{N} - \frac{(a+c)(a+d)}{N^2}$$

$$= \tfrac{1}{4}(1+\alpha_2)(1+\alpha_1) - \tfrac{1}{4}(1+\alpha_1)(1+\alpha_1\alpha_2)$$

$$= \tfrac{1}{4}\alpha_2(1-\alpha_1^2),$$

and similarly: $$\frac{ac-bd}{N^2} = \tfrac{1}{4}\alpha_1(1-\alpha_2^2).$$

Hence substituting in (i)

$$\text{Probable error of } r = \frac{\cdot 67449}{\sqrt{NHK}} \{\tfrac{1}{16}(1-\alpha_1^2\alpha_2^2) + \tfrac{1}{16}\alpha_2^2(1-\alpha_1^2) + \tfrac{1}{16}\alpha_1^2(1-\alpha_2^2) - \tfrac{1}{8}\alpha_2^2(1-\alpha_1^2) - \tfrac{1}{8}\alpha_1^2(1-\alpha_2^2)\}^{\frac{1}{2}}$$

$$= \frac{\cdot 67449}{\sqrt{NHK}} \sqrt{\tfrac{1}{16}(1-\alpha_1^2)(1-\alpha_2^2)} \quad \text{.............................(iii).}$$

This can also be put in the form:

$$\text{Probable error of } r = \frac{\cdot 67449}{\sqrt{NHK}} \sqrt{\frac{(a+b)(a+c)(d+b)(d+c)}{N^4}} \quad \text{......(iv).}$$

This is the probable error of r of a fourfold table when the real value of r is zero.

Now as (ii) and (iv) give the reducing factors for the two cases (a) when h and k are both zero but r has any value and (b) when h and k have any values but r is zero, it occurred to me that the combined product of the two would give good results for a considerable range of values of h and k and r. We have to note that (iv) for h and k zero becomes

$$\frac{\cdot 67449}{\sqrt{N}} \frac{\pi}{2}.$$

Hence we take as our formula:

Probable error of r

$$= \frac{\cdot 67449}{\sqrt{N}} \sqrt{1-r^2} \sqrt{1 - \left(\frac{\sin^{-1} r}{90^\circ}\right)^2} \frac{\sqrt{\frac{1}{2}(1+\alpha_1)\frac{1}{2}(1-\alpha_1)}}{H} \frac{\sqrt{\frac{1}{2}(1+\alpha_2)\frac{1}{2}(1-\alpha_2)}}{K} \quad \text{...(v).}$$

Now it will be seen that this consists of three parts:

(a) $\sqrt{1-r^2}\sqrt{1-\left(\frac{\sin^{-1} r}{90°}\right)^2}$. This is easy to table for all values of r.

(b) $\frac{\sqrt{\frac{1}{2}(1+\alpha_1)\frac{1}{2}(1-\alpha_1)}}{H}$, and

(c) $\frac{\sqrt{\frac{1}{2}(1+\alpha_2)\frac{1}{2}(1-\alpha_2)}}{K}$.

Both these (b) and (c) can be readily found from a single table rapidly formed from Sheppard's Table of the Probability Integral. The entry to the single table will be $(a+c)/N$ or $(a+b)/N$, i.e. $\frac{1}{2}(1+\alpha)$.

Thus a knowledge of the correlation r and the two division percentages (together with Miss Gibson's Table for $\cdot 67449/\sqrt{N}$), will enable us by the aid of the two new tables to rapidly write down four factors whose product gives the required probable error. I have tested the form (v) against the true probable error as found from (i). In all cases it gave results differing only from the true value at most by about one or two units in the third place of figures—a result amply accurate for all practical purposes.

Illustration I.

211·25	153·75	365
152·75	560·25	713
364	714	1078

The correlation was found to be ·5557 ± ·0261; the probable error from the short formula was ·0265.

Illustration II.

1562	42	1604
383	94	477
1945	136	2081

The correlation was found to be ·5954 ± ·0272; the probable error from the short formula was ·0293.

Illustration III.

455	622	1077
599	1324	1923
1054	1946	3000

The correlation was found to be ·1811 ± ·0210; the probable error from the short formula was ·0199.

Illustration IV.

849	665	1514
205	1281	1486
1054	1946	3000

The correlation was found to be ·6633 ± ·0132; the probable error from the short formula was ·0132.

Illustration V.

1196	223	1419
318	1263	1581
1514	1486	3000

The correlation was found to be ·8464 ± ·0079; the probable error from the short formula was ·0079.

These examples will suffice, I think, to give confidence in the formula and in the tables accompanying this paper. The absence of probable errors from the expressions for fourfold table correlations can no longer be justified on the ground of their great laboriousness.

The following Tables have been calculated by Miss Julia Bell.

Let $\chi_1 = \cdot 67449/\sqrt{N}$. This is given by Miss Gibson's Tables, *Biometrika*, Vol. III. p. 387. Let

$$\chi_r = \sqrt{1-r^2}\sqrt{1-\left(\frac{\sin^{-1} r}{90^\circ}\right)^2},$$

and

$$\chi_\alpha = \frac{1}{H}\sqrt{\tfrac{1}{2}(1+\alpha)\times\tfrac{1}{2}(1-\alpha)}.$$

Then

$$\text{Probable error of } r = \chi_1 \cdot \chi_r \cdot \chi_{\alpha_1} \cdot \chi_{\alpha_2}.$$

TABLE I. *Values of χ_r for Values of r.*

r	χ_r	r	χ_r	r	χ_r	r	χ_r	r	χ_r
·00	1·0000	·20	·9717	·40	·8845	·60	·7298	·80	·4843
·01	·9999	·21	·9688	·41	·8785	·61	·7200	·81	·4687
·02	·9997	·22	·9657	·42	·8723	·62	·7099	·82	·4526
·03	·9994	·23	·9625	·43	·8659	·63	·6997	·83	·4362
·04	·9989	·24	·9591	·44	·8594	·64	·6892	·84	·4192
·05	·9982	·25	·9556	·45	·8527	·65	·6785	·85	·4018
·06	·9975	·26	·9520	·46	·8458	·66	·6675	·86	·3838
·07	·9966	·27	·9482	·47	·8388	·67	·6563	·87	·3652
·08	·9955	·28	·9442	·48	·8315	·68	·6448	·88	·3461
·09	·9943	·29	·9401	·49	·8241	·69	·6331	·89	·3262
·10	·9930	·30	·9358	·50	·8165	·70	·6211	·90	·3057
·11	·9915	·31	·9314	·51	·8087	·71	·6088	·91	·2843
·12	·9899	·32	·9268	·52	·8007	·72	·5962	·92	·2620
·13	·9881	·33	·9221	·53	·7926	·73	·5834	·93	·2387
·14	·9862	·34	·9172	·54	·7842	·74	·5702	·94	·2142
·15	·9841	·35	·9122	·55	·7756	·75	·5568	·95	·1882
·16	·9819	·36	·9070	·56	·7669	·76	·5430	·96	·1605
·17	·9796	·37	·9016	·57	·7579	·77	·5288	·97	·1305
·18	·9771	·38	·8961	·58	·7488	·78	·5144	·98	·0972
·19	·9745	·39	·8904	·59	·7394	·79	·4995	·99	·0585
								1·00	·0000

TABLE II.

Values of χ_α for Values of $\frac{1}{2}(1+\alpha)$.

$\frac{1}{2}(1+\alpha)$	χ_α	$\frac{1}{2}(1+\alpha)$	χ_α	$\frac{1}{2}(1+\alpha)$	χ_α	$\frac{1}{2}(1+\alpha)$	χ_α
·50	1·2533	·65	1·2877	·80	1·4288	·95	2·1132
·51	1·2535	·66	1·2928	·81	1·4457	·96	2·2740
·52	1·2539	·67	1·2984	·82	1·4641	·97	2·5071
·53	1·2546	·68	1·3044	·83	1·4844	·98	2·8915
·54	1·2556	·69	1·3109	·84	1·5067	·985	3·2097
·55	1·2569	·70	1·3180	·85	1·5315	·990	3·7333
·56	1·2585	·71	1·3256	·86	1·5590	·991	3·8854
·57	1·2604	·72	1·3338	·87	1·5897	·992	4·0639
·58	1·2626	·73	1·3427	·88	1·6245	·993	4·2784
·59	1·2652	·74	1·3523	·89	1·6640	·994	4·5419
·60	1·2680	·75	1·3626	·90	1·7094	·995	4·8779
·61	1·2712	·76	1·3738	·91	1·7623	·996	5·3278
·62	1·2748	·77	1·3859	·92	1·8249	·997	5·9776
·63	1·2787	·78	1·3990	·93	1·9003	·998	7·0465
·64	1·2830	·79	1·4133	·94	1·9937	·999	9·3870

MULTIPLE CASES OF DISEASE IN THE SAME HOUSE.

APPENDIX TO PAPERS IN *BIOMETRIKA*, VOL. VIII. p. 404 AND p. 430.

By KARL PEARSON, F.R.S.

I REGRET that a most careless algebraic slip has crept into my work on this subject. It stared me in the face when I saw the published number of *Biometrika*, and I cannot understand how it escaped me in MS. or proof.

After obtaining the fundamental equations

$$\sigma_{p_t}\sigma_{p_s}R_{p_tp_s} = -\frac{st\bar{p}_s\bar{p}_t}{N} \quad \text{....(xvi)},$$

$$\sigma^2_{p_s} = \bar{p}_s\left(1 - \frac{\bar{p}_s s^2}{N}\right) \quad \text{....(xvii)},$$

where

$$N = n\left(1 + \frac{n}{m}\right) \quad \text{....(xviii)},$$

on p. 410, I continued on p. 411:

"Now let us write $x_s = s^2 p_s$; then clearly $\sigma^2_{x_s} = s^2\sigma^2_{p_s}$ ...." Clearly nothing is more false! $\sigma^2_{x_s}$ would then equal $s^4\sigma^2_{p_s}$ and not $s^2\sigma^2_{p_s}$, and thus the value of χ^2, i.e.

$$\overset{u}{\underset{1}{S}}\left\{\frac{s^2(p_s - \bar{p}_s)^2}{\bar{p}_s}\right\},$$

is erroneous.

It is not possible I think to obtain the value of χ^2 in this indirect and brief method. We must return to the multiple correlation formulae of my memoir in the *Phil. Mag.*, July, 1900, p. 161, and evaluate the determinant R and its minors for the special values (xvi) and (xvii).

Following the lines of that paper let us write:

$$\bar{p}_s s^2/N = \sin^2\beta_s \quad \text{....(xix)},$$

then

$$\sigma^2_{p_s} = \frac{N\cos^2\beta_s \sin^2\beta_s}{s^2},$$

and

$$\sigma_{p_s} = \sqrt{N}\cos\beta_s \sin\beta_s/s \quad \text{....(xx)}.$$

Hence from (xvi):

$$\frac{N\cos\beta_t\sin\beta_t}{t}\,\frac{\cos\beta_s\sin\beta_s}{s}\,R_{p_tp_s} = -\frac{stN^2\sin^2\beta_s\sin^2\beta_t}{s^2t^2N},$$

or

$$R_{p_tp_s} = -\tan\beta_s\tan\beta_t \;\;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxi)},$$

and agrees entirely with the value of r_{ts} used in my *Phil. Mag.* paper just referred to.

In other words the determinant R and its minors R_{ss} and R_{st} will take precisely the same forms as in that paper provided we

(i) replace Eq. (xi) of that paper by our result from (xix) above, i.e.

$$\eta_s = \cot^2\beta_s = \frac{N}{\bar{p}_s s^2} - 1 \;\;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxii)},$$

and

(ii) remember that we must reduce the total number of our variants

$$p_0,\, p_1,\, \ldots\, p_u$$

by the two relations $\qquad S(p_s) = m, \qquad S(sp_s) = n \;\;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxiii)},$

where m is the total number of houses and n of cases. That is to say our total number of frequency groups, which are not wholly dependent but are correlated, is

$$u + 1 - 2 = u - 1.$$

We select as our dependent variates, which are fixed as soon as the others are known, p_0 and p_u. Thus the value of the auxiliary determinant J, p. 162 of my *Phil. Mag.* memoir, is:

$$J = \begin{vmatrix} -\eta_1 & 1 & 1 & \ldots & 1 \\ 1 & -\eta_2 & 1 & \ldots & 1 \\ 1 & 1 & -\eta_3 & \ldots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \ldots & -\eta_{u-1} \end{vmatrix}.$$

If $\qquad \lambda = (1+\eta_1)(1+\eta_2)(1+\eta_3)\ldots(1+\eta_{u-1}),$

we find as in that memoir:

$$J = (-1)^n\lambda\left(1 - \frac{1}{1+\eta_1} - \frac{1}{1+\eta_2} - \ldots - \frac{1}{1+\eta_{u-1}}\right) \;\;\ldots\ldots\ldots\text{(xxiv)},$$

$$J_{st} = (-1)^{n-1}\frac{\lambda}{(1+\eta_s)(1+\eta_t)} \;\;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxv)},$$

$$J_{ss} = (-1)^{n-1}\frac{\lambda}{1+\eta_s}\left\{1 - \frac{1}{1+\eta_1} - \frac{1}{1+\eta_2} - \ldots \right.$$

$$\left. - \frac{1}{1+\eta_{s-1}} - \frac{1}{1+\eta_{s+1}} - \ldots - \frac{1}{1+\eta_{u-1}}\right\} \;\;\ldots\ldots\text{(xxvi)}.$$

To evaluate these we note that

$$\overset{u-1}{\underset{1}{S}}\left(\frac{1}{1+\eta_s}\right)=\overset{u-1}{\underset{1}{S}}(\sin^2\beta_s)=\overset{u-1}{\underset{1}{S}}\frac{(\bar{p}_s s^2)}{N},$$

but $\overset{u}{\underset{1}{S}}(\bar{p}_s s^2)=N$ as shewn on p. 411 of my *Biometrika* memoir. Hence:

$$\overset{u-1}{\underset{1}{S}}\left(\frac{1}{1+\eta_s}\right)=\frac{1}{N}(N-\bar{p}_u u^2).$$

Thus we have:

$$J=(-1)^n\lambda\frac{\bar{p}_u u^2}{N},$$

$$J_{ss}=\frac{(-1)^{n-1}\lambda}{1+\eta_s}\left(\frac{\bar{p}_s s^2}{N}+\frac{\bar{p}_u u^2}{N}\right),$$

$$J_{st}=\frac{(-1)^{n-1}\lambda}{(1+\eta_s)(1+\eta_t)}.$$

Next as in the *Phil. Mag.* paper (p. 161):

$$R=(-1)^n J/\lambda,$$

$$R_{ss}=(-1)^{n-1}J_{ss}\cot^2\beta_s/\lambda,$$

$$R_{st}=(-1)^{n-1}J_{st}\cot\beta_s\cot\beta_t/\lambda.$$

Thus:

$$R=\bar{p}_u u^2/N,$$

$$R_{ss}=\cos^2\beta_s(\bar{p}_s s^2+\bar{p}_u u^2)/N,$$

$$R_{st}=\cos\beta_s\sin\beta_s\cos\beta_t\sin\beta_t.$$

Finally we have:

$$\frac{R_{ss}}{R\sigma^2_{p_s}}=\frac{\cos^2\beta_s(\bar{p}_s s^2+\bar{p}_u u^2)}{\bar{p}_u u^2}\times\frac{1}{\cos^2\beta_s\bar{p}_s}$$

$$=\frac{1}{u^2}\left(\frac{s^2}{\bar{p}_u}+\frac{u^2}{\bar{p}_s}\right)=\frac{s^2}{\bar{p}_u u^2}+\frac{1}{\bar{p}_s}\ldots,$$

and

$$\frac{R_{st}}{R\sigma_{p_s}\sigma_{p_t}}=\frac{\cos\beta_s\sin\beta_s\cos\beta_t\sin\beta_t}{\bar{p}_u u^2/N}\times\frac{st}{N\cos\beta_s\sin\beta_s\cos\beta_t\sin\beta_t}$$

$$=\frac{1}{u^2}\frac{st}{\bar{p}_u}.$$

Thus finally χ^2 is given by

$$\overset{u-1}{\underset{1}{S}}\frac{R_{ss}}{R\sigma^2_{p_s}}(p_s-\bar{p}_s)^2+2S\frac{R_{st}}{R\sigma_{p_s}\sigma_{p_t}}(p_s-\bar{p}_s)(p_t-\bar{p}_t)$$

$$=\overset{u-1}{\underset{1}{S}}\left\{\left(\frac{s^2}{\bar{p}_u u^2}+\frac{1}{\bar{p}_s}\right)(p_s-\bar{p}_s)^2\right\}+2S\left\{\frac{st}{u^2\bar{p}_u}(p_s-\bar{p}_s)(p_t-\bar{p}_t)\right\}$$

$$=\overset{u-1}{\underset{1}{S}}\frac{(p_s-\bar{p}_s)^2}{\bar{p}_s}+\frac{1}{u^2\bar{p}_u}\left\{\overset{u-1}{\underset{1}{S}}s(p_s-\bar{p}_s)\right\}^2.$$

But

$$\overset{u}{\underset{1}{S}}s(p_s-\bar{p}_s)=0,$$

therefore

$$\overset{u-1}{\underset{1}{S}}s(p_s-\bar{p}_s)=-u(p_u-\bar{p}_u).$$

Thus we have: $$\chi^2 = \overset{u}{\underset{1}{S}} \left\{ \frac{(p_s - \bar{p}_s)^2}{\bar{p}_s} \right\}.$$

In other words χ^2 is to be found from the ordinary χ^2 for the $u+1$ variates *excluding the houses with zero cases.* I had excluded these houses from my weighting formula, i.e.

$$\chi^2 = \overset{u}{\underset{1}{S}} \left\{ \frac{s^2 (p_s - \bar{p}_s)^2}{\bar{p}_s} \right\},$$

which was, however, in error. The higher multiple cases are now we see *not* heavily weighted, but we are not to use the frequency of the zero case houses in evaluating our χ^2. This is the effect of the double relation between the $u+1$ p's. An additional p is cast out of our result as compared with the ordinary frequency problem of goodness of fit.

I have next to consider how far this correction modifies the results obtained in my papers for numerical cases.

For enteric cases we have as on p. 412 of my *Biometrika* paper

$$\chi^2 = \frac{(3398 - 3350)^2}{3398} + \frac{(78 - 56)^2}{56} + \frac{(2 - 1)^2}{1}$$
$$= \cdot 678 + 8\cdot 643 + 1 = 10\cdot 321,$$

hence we have for $n' = 3$, $P = \cdot 006$, or the odds are about 166 to 1 against such a large divergence from chance.

For cancer cases we have from the returns on p. 432

$$\chi^2 = \frac{(2\cdot 6)^2}{312\cdot 4} + \frac{(9\cdot 4)^2}{20} + \frac{(4\cdot 16)^2}{1\cdot 84} + \frac{(\cdot 914)^2}{\cdot 086}$$
$$= \cdot 022 + 4\cdot 418 + 9\cdot 405 + 9\cdot 714 = 23\cdot 559.$$

For $n' = 4$, we have $P = \cdot 0003$ or the cancer house distribution is a very improbable one.

If we deal with the experimental data that were obtained for probabilities on the same basis as the cancer statistics we have:

$$\chi^2 = \frac{3^2}{313} + \frac{3^2}{29} + \frac{1}{2}$$
$$= \cdot 029 + \cdot 310 + \cdot 500 = \cdot 839.$$

Or, for $n' = 3$, the probability is over $\cdot 60$.

Thus although I made a bad algebraic slip the new values confirm practically the old and Dr Webb's cancer data suggest that there may very possibly exist a relationship between cancer and environment of some kind.

Since my papers in *Biometrika* were written, my attention has been drawn to a *Special Report on Cancer in Ireland* which was issued in 1903 as a supplement

to the 38th *Annual Report* of the Registrar-General for Ireland. On p. 34 we find that in the ten years 1876–1885 inclusive there were 12 multiple cancer houses in the City of Dublin. In all these instances in only one house did the two deaths occur in members of the same family. There are no cases—or at least no cases recorded—of more than two deaths in the same house.

Unfortunately the paper does not give either (i) the number of inhabited houses in the City of Dublin during the ten years under consideration, or (ii) the total number of deaths in the City of Dublin from cancer. What useful purpose the then Registrar-General could conceive such data would serve completely puzzles me*. He draws up a summary of his Tables, introducing it with the words: "I venture to draw attention to some of the main facts which they disclose," and clause (7) runs:

"That in some instances more than one case of cancer has occurred amongst different families living in the same house, or amongst successive occupants of the same house."

Now unless this is meant to be interpreted as a suggestion that the multiple houses are in excess, it must be anticipated from the mere random occurrence of cancer. Anyhow without further information the data on this point, as on many others in this Special Report on Cancer in Ireland, are wholly worthless and the publication does no credit to a Government Department.

I have striven to obtain the requisite additional data from the present Registrar-General for Ireland. He most kindly informs me that the inhabited houses of the *City* of Dublin numbered 23,896 in 1871 and 24,211 in 1881; and for the Registration Area of Dublin, 34,118 in 1871 and 36,232 in 1881. All these data are from the Censuses of those years. The deaths from cancer in the Registration Area for 1876 to 1885 inclusive were 1714, but how many of these occurred in the City of Dublin he is not able to tell me. I presume the late Registrar-General must have known these deaths in order to detect multiple houses, but apparently they cannot now be ascertained. It is clear therefore that the data of the "Special Report on Cancer" must remain practically worthless.

If we suppose the City of Dublin to have had a number of cancer cases proportionate to its houses we might take:

$$\text{Cancer Cases in City} = \frac{\frac{1}{2}(23{,}896 + 24{,}211)}{\frac{1}{2}(34{,}118 + 36{,}232)}1714 = 1172,$$

roughly $\frac{2}{3}$ of the cases in the Registration Area.

Since the middle of our period is not very far from the Census year 1881, we might take

$$n = 1172, \qquad m = 24{,}211.$$

* Cancer in the Report is associated with alcohol, syphilis, smoking, etc., in an equally unscientific manner. Unless we know the incidence of each character or disease in the population, how is it possible to determine whether the association is merely due to chance or not?

Whence we deduce:

Calculated	Observed
$\bar{p}_1 = 1116{\cdot}6$	$\bar{p}_1 = 1148$
$\bar{p}_2 = 27{\cdot}0$	$\bar{p}_2 = 12$
$\bar{p}_3 = 0{\cdot}4$	$\bar{p}_3 = 0$

$\chi^2 = 8{\cdot}7$ and $P = {\cdot}013$, i.e. odds of about 76 to 1.

Here our deviations are not as before towards multiple houses, but *cancer apparently avoids a house where it has paid one visit!* It is difficult to believe that the frequency should be so far below a random distribution in the case of multiple houses, and the matter would be still worse, if there were more cases relative to the houses in the City than in the Registration Area. I think we may safely conclude that the data for multiple houses provided by the Registrar-General for Ireland are probably more than 100 per cent. in error, and therefore are wholly worthless for the purposes for which they are apparently stated. We cannot suppose one visit of cancer to confer immunity on a house, and the fact that the multiple houses are so significantly short of the chance frequency suffices, I think, to discredit the data and the accuracy of the methods adopted in this Special Report on Cancer.

A STUDY OF THE VARIATIONS IN THE FEMALE PELVIS, BASED ON OBSERVATIONS MADE ON 217 SPECIMENS OF THE AMERICAN INDIAN SQUAW*.

By ARTHUR BREWSTER EMMONS, A.B., M.D. (Harvard), Boston, Mass., U.S.A.

THE object of this investigation was to determine, so far as possible, the variation in form of the "normal" human female pelvis. By "normal" it was intended to exclude all pathological pelves, and to include all variations, not the results of disease. Our conception of "normal," for a standard of comparison, should, I believe, include not only the average measurements, but also the minimum and maximum as well as the proportion of cases at regular intervals between these extremes.

The bones of the American Indian of the earlier times, collected from various parts of North and South America and the adjacent islands, are said by the authorities on the Indian (1) to be entirely free from rhachitis, and that other diseases affecting the bones, as tuberculosis, osteomalacia, and syphilis, are rare. This statement was borne out in my series of specimens, for no evidence of these diseases was encountered. In an occasionally elderly specimen, however, the remains of an old osteo-arthritic process were found. This late change could have had little or no effect on the form of the pelves.

Varying conditions, exclusive of disease, such as differences of nutrition, and certain habits as sitting up for long periods in early infancy, carrying heavy burdens in youth, are the common causes and most probable influences modifying the shape of the pelvis. Among the many Indian tribes from which the specimens of this series came, these factors may have been present at times in varying degrees, yet it seems fair to consider them all as from a pure unmixed race, and thus the series should yield a true type.

A due proportion of variations can scarcely be obtained with certainty unless at least 200 specimens are used. A still larger series than mine might yield

* Awarded the Boylston Medical Prize Essay, 1912. "The Boylston Medical Committee do not consider themselves as approving the doctrines contained in any of the dissertations to which premiums may be adjudged."

slight or occasional variations from the standard here set. But it is felt that this series of 217 specimens will form a group sufficiently comprehensive to include all important variations in approximately their due proportions, and serve for a standard of comparison with the male pelvis and the pelves of animals and other races of man, as well as a standard to differentiate the pathological.

Material. The material on which this paper is based is to be found largely in the splendid collection of the U.S. National Museum in Washington. The remainder is in the Peabody Museum, Cambridge, Mass., and the American Natural History Museum, New York City. To the authorities of these institutions the writer wishes to express his thanks for the privilege of using the specimens, which has made this study possible. He also desires to extend his grateful acknowledgments to Dr Hrdlicka of the U.S. National Museum for his kindness and advice.

To obtain accuracy and to avoid error several special means were employed. A few of the specimens were somewhat injured. The exact relative position of these bones could not always be accurately fixed. In a few cases careful estimates only could be made, but no specimens were used unless the relative positions of the bones could be obtained with a fair degree of accuracy. When broken specimens were used the measurement was marked ("a") approximate. The use of the accurate "compas glissière" for the greater part of the measurements also tended to reduce error. Nearly all measurements were made on disarticulated pelves. To hold the bones in proper relative positions the hand and the sand-box were found rather unsatisfactory. The apparatus seen in Plate V was therefore devised and used for the entire series. It is believed that the margin of unavoidable error in these measurements is not a great one.

Measurements. Anthropologists differ widely in what they have considered the essential measurements of the pelvis. These differences probably arose because the pelvis is such a complicated architectural bony structure.

Topinard (2) gave measurements on 207 pelves of animals and man. Verneau (3) made a large number of measurements on specimens from many different races. Turner (4), reducing the number of measurements made by Verneau, gave observations on specimens of many peoples collected on the *Challenger* expedition. From these I have selected only those few measurements which seemed to me to be most essential for comparison, and I have added a very few observations of interest in regard to one particular region, the pelvic outlet.

In the last step in the evolution of man either from the primate to man or from one class of primate to another, whichever anthropological classification of man is used, the most essential change which concerned the pelvis was the assumption of the erect posture. The pelvis has carried on the function of child-bearing probably ever since a pelvis has existed. The function of weight-bearing, however, by the recent assumption of the upright attitude was transferred

from four legs to the hind two. The pelvis, now become a very important girder in the bridge structure, thus came to bear the weight not only of the fore part of the body but also of the head, and this head has since increased considerably in weight. With this large increase in the weight to be carried by the pelvis has come a marked change in the direction in which that weight is applied. These three factors, the change in direction, the extra weight, and the shortness in time since this change has taken place, would lead us to expect changes of form of considerable importance and with them variations, the frequent accompaniment of recent change, in those parts to which this function applies.

The typical "male" pelvis, free from the function of child-bearing, is built strong, high, close-knit, thick-boned, with a small cavity, the type best adapted to bearing weight. The typical "female" pelvis is, on the other hand, of lighter build, lower, more open, and with a roomy cavity, better adapted to child-bearing. The newly-acquired function of weight-bearing, however, tends to mould the female pelvis toward the male type, while the child-bearing function resists such changes in so far as they tend to interfere with its long-established and most vital function of child-bearing. To see how nature arbitrates between these opposing forces, in other words how with this change the function of child-bearing is preserved, was one of the reasons for selecting certain measurements. I have also included those diameters and indices with which to compare various groups and races.

Modern pelvimetry, as employed in obstetrics, regards the "obstetric conjugate,"* the shortest distance between promontory and pubic symphysis, as the chief factor of pelvic efficiency. Secondary in importance to this diameter and correlated with it, is the breadth of the pelvic "inlet," as shown by the greatest transverse diameter. A comparison of these two dimensions indicates roughly the shape of this space. Lastly, it has been recently emphasized that the pelvic "outlet" is a not infrequent cause of difficulty in labour and, if contracted, its efficiency may be gauged by the inter-tuberal diameter correlated with the space posterior to it. The more exact estimate of these two spaces, so essential to child-birth, and the means of calculating them in the living, are the direct objects of certain of the measurements taken.

Inlet. In life direct mensuration of the "obstetric conjugate" (Diagram I, *A—R*, and Plate II) or the conjugata vera cannot be satisfactorily performed. Calculations in life from the "external conjugate," the distance from the tip of the spinous process of the last lumbar vertebra behind to the top of the pubic symphysis in front, as introduced by Baudeloque (5), though often suggestive are admittedly unreliable. We have, however, a much more reliable measure for calculation of this vera, which is the oblique or diagonal conjugate (Diagram I, *A—P*), extending from the promontory of the sacrum to the lower border of the symphysis, as

* The "conjugata vera" of the text-books is taken from the promontory of the sacrum to the top of the pubic symphysis and is a variable amount longer than the "obstetric conjugate," consequently I prefer this shorter measure though both terms are used synonymously.

measured in life without difficulty in small pelves by means of two fingers through the vagina. To calculate the vera after obtaining the diagonal conjugate the inner surface of the pubic symphysis is palpated and two points observed; first, the height of the point nearest the promontory above the symphysis (Diagram I, *R—P*); second, the angle which the line on which these two points lie makes with the diagonal conjugate (Diagram I, *R—P—A*). From these data we may calculate in each case the amount to be subtracted from the diagonal conjugate to give us the obstetric conjugate or vera (Diagram I, $AP - NP = AR$). This difference (*NP*) was calculated in the dried specimens, and found to vary from 0·8 cm. to 3·2 cms. in my series.

The following table gives the data used in the living for estimating the "obstetric diameter" or the conjugata vera. The pubic height was taken from the point on the pubic symphysis nearest the promontory, as an upper limit, to the point on the pubic symphysis nearest the tip of the sacrum, as the lower limit, which space may be called the length of pubic resistance (see Plate VI, 6 and Diagram I, *R—P*).

Oblique Diameter	Freq.	Obstetric Diameter	Freq.	Difference	Freq.	Pubic Height	Freq.
9·4 to 9·9 cms.	4	7·5 to 8·9 cms.	14	0·8 to 1·4 cms.	46	1·1 to 1·9 cms.	9
10 to 10·9 „	17	9 to 9·9 „	60	1·5 to 1·9 „	98	2 to 2·9 „	109
11 to 11·9 „	74	10 to 10·9 „	75	2 to 2·4 „	58	3 to 4·3 „	95
12 to 12·9 „	73	11 to 11·9 „	54	2·5 to 3·2 „	14	broken	4
13 to 13·9 „	38	12 to 12·9 „	10	—	—	—	—
14 to 14·9 „	10	13 to 14 „	4	—	—	—	—
Average 11·78 cms.		10·68 cms.		1·76 cms.		2·8 cms.	
Maximum 14·9 „		14 „		3·2 „		4·3 „	
Minimum 9·4 „		7·5 „		0·8 „		1·1 „	

The antero-posterior distance being thus estimated, we turn to the measure of the width of the inlet, the transverse diameter. This distance cannot be satisfactorily measured in the living. Estimates of this diameter made in the usual way from the maximum inter-cristal and inter-spinous diameters are suggestive but unreliable, as demonstrated by Scheffer (6), who reported a difference of 3·3 cms. in the inter-cristal measures of two pelves with equal transverse diameters. My series contains two rather narrow pelves (♯62 and ♯200), the comparison of which bears on this point; one a generally contracted or "pygmy" pelvis, the other simply narrow. Three pelves, moreover, with nearly equally broad inlets, show wide variations in their inter-cristal diameters:—

♯62 crests 20·8 cms.	spines 17·2	transverse 10·7	obstetric diameter 8·5
♯200 „ 23 „	„ 19·5	„ 10·3	„ „ 11
♯30 „ 25·5 „	„ 23·25	„ 14	„ „ 9·8
♯112 „ 29·1 „	„ 25	„ 14·1	„ „ 12
♯143 „ 25·5 „	„ 22	„ 14	„ „ 9·2

It may be seen from this table that #200 with inter-cristal diameter over 2 cms. broader than #62, yet has a narrower transverse diameter. Again #112, with a transverse diameter practically identical with #30 and #143, has an inter-cristal diameter 3·6 cms. greater. By these extreme differences we are forced to conclude that accurate estimates of the transverse diameter of the inlet cannot be deduced from inter-cristal and inter-spinous diameters.

Since we cannot measure satisfactorily the transverse diameter in the living, we are therefore forced to adopt for obstetric use some such rough rule as Williams (7) gives as follows:—"Despite many inaccuracies, the external measurements are of considerable value, in that they serve to indicate with tolerable certainty the variety of pelves (contracted) with which one has to deal. Normally the distance between the spines is 2·5 to 3 cms. less than between the crests: but in rhachitic pelves, owing to the flaring of the iliac bones, this proportion becomes deranged, and the two measurements approximate one another in length, the former frequently being equal to, and occasionally exceeding, the latter. If, however, both measurements are considerably below the normal, but preserve their usual relation to one another, and at the same time the external conjugate is also shortened proportionately, it is permissible to conclude that the entire pelvis measures below normal in all its diameters, or, in other words, is generally contracted." The following summarising table of these measures may make the results of this series more graphic:—

Inter-cristal	Freq.	Inter-spinous		Freq.	Transverse of Inlet	Freq.
20·8 to 21·9 cms.	2	17·2 to 18·9 cms.		2	10·3 to 10·9 cms.	2
22 to 23·4 „	10	19 to 20·9 „		28	11 to 11·9 „	14
23·5 to 24·9 „	42	21 to 22·4 „		67	12 to 12·9 „	71
25 to 26·4 „	88	22·5 to 23·9 „		67	13 to 13·9 „	108
26·5 to 27·9 „	59	24 to 25·4 „		44	14 to 14·7 „	22
28 to 29·1 „	16	25·5 to 27 „		9	—	—
Average 25·76 cms.		22·66 cms.	difference 3·1 cms.		12·95 cms.	
Maximum 29·1 „		27 „	„ 6·75 „		14·7 „	
Minimum 20·8 „		17·2 „	„ 1 „		10·3 „	

While the average difference between the inter-cristal and the inter-spinous was 3·1 cms., the variation between the extremes (1 to 6·75 cms.) of this series free from disease is wide and would tend to indicate that little reliance could be placed on deductions from such a comparison.

$$\textit{Superior strait Index} = \frac{\text{obstetric conjugate (vera)} \times 100}{\text{transverse diameter}},$$

average 79·5, maximum 107·7, minimum 61·5.

These data were considered sufficient to give as accurate a measure of the inlet

as possible. For simplification the oblique diameters of the inlet, which cannot be measured on the living, were not included.

Outlet. The pelvic outlet is commonly of very secondary importance as compared to the inlet. The space is a complicated one of irregular outline, and has received much less attention for these reasons. Occasional extreme contractions of this space have been recorded, usually, however, only after disastrous results of labour have occurred. The subject has lately been thoroughly reviewed historically from an obstetrical standpoint by Williams (8), who gives also his clinical observations on 1200 women. In analyzing this complex space a few special measurements were taken in the hope of rendering more simple and more exact the observations necessary for its estimation (see Plates III, VI, 7, VII, 8, 9, 10, and Diagrams I, II, and III).

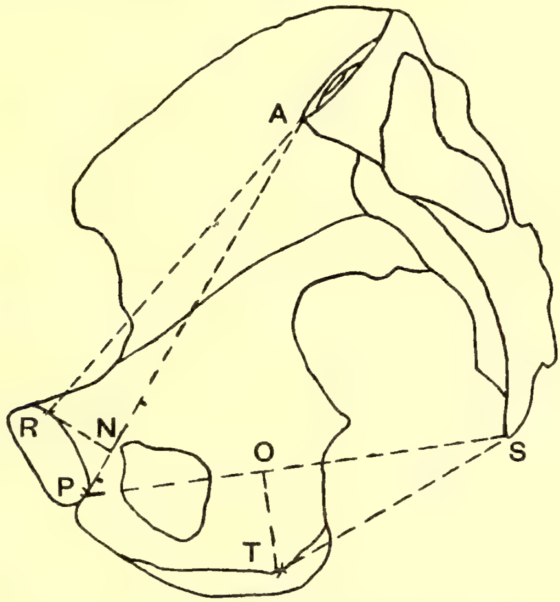


Diagram I (traced from Plate VI, 7, pelvis #106).

Diameters of the Inlet. Planes of the Outlet.

A.	Promontory.
R.	Nearest point of pubic symphysis to promontory.
A—R.	Inlet, obstetric diameter, or vera.
P.	Nearest point of pubic symphysis to sacrum.
R—P.	"Height of Pubic Resistance."
A—P.	Diagonal conjugate.
N—R.	Perpendicular from diagonal conjugate to obstetric conjugate.
N—P.	Amount subtracted from diagonal conjugate to obtain obstetric conjugate.
S.	Tip of sacrum.
T.	Tuberosity, point of impingement.
S—P.	Antero-posterior diameter of the outlet.
O—T.	Perpendicular from antero-posterior diameter of outlet to inter-tuberal diameter.
S—T.	Posterior sagittal diameter.
P—O.	Pubic symphysis to foot of perpendicular.

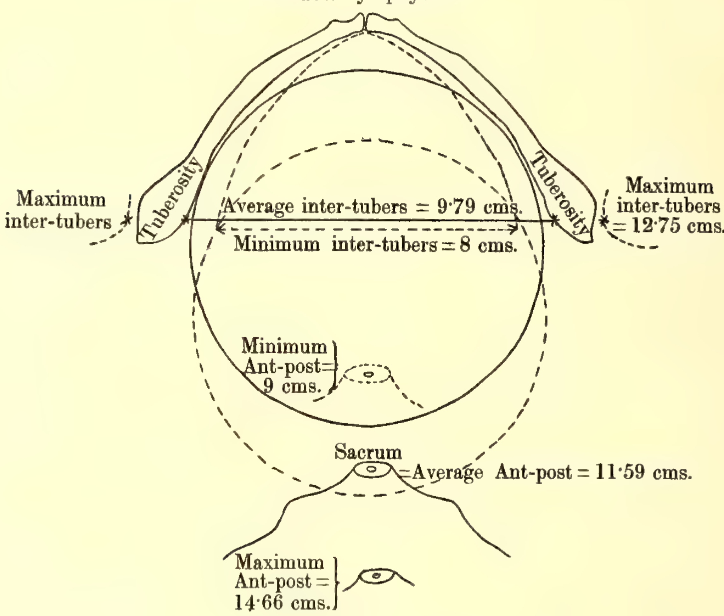


DIAGRAM II. Outlet.

Circles represent the average fetal head circumference, 9·5 cms.
Average inter-tuberal diameter ———— .
Minimum ,, ,, - - - - - - .
Maximum ,, ,, indicated.

This diagram shows the effect of the narrowing of the inter-tuberal diameter.

A contracted outlet, as is clearly pictured in Williams' article, means a narrowing of the space between the tuberosities, accompanied usually by the so-called "male arch," combined with a short antero-posterior diameter. The obstruction caused by this narrowing may be greatly increased by a forward position of the tip of the sacrum, or may be decreased by a backward position of that prominence (see Diagram III). The determination of the inter-tuberal diameter is, therefore, of prime importance, that of the "posterior-sagittal" (Diagram I, *S*—*T*), the distance from the tip of the sacrum to the inter-tuberal line, is secondary to it and correlated with it.

One of the first difficulties met with in estimating the outlet was the determination of the points on the tuberosities from which to measure. Owing to the soft parts this difficulty is even greater in the living. While in some pelves the angle of the tuberosities was such that the points could easily be selected, in others the rounded and complex curves made this selection a mere guess. To overcome this difficulty and to obtain consistent results, a circular sheet of transparent celluloid was marked with concentric circles 0·5 centimetre apart (Plate VII, 8). This sheet was applied to the inner edge of the tip of the sacrum, as one point, and to the two tuberosities as the other essential points

of resistance (cf. Diagrams II and III). The circle which just touched these three points was read off as representing the size of the fetal head which might pass that particular outlet, and the points of contact were marked as the "tuberosities." Often the circle was tangent for a short distance to the tuberosities. The middle of this line of tangency was taken as the point of maximum resistance.

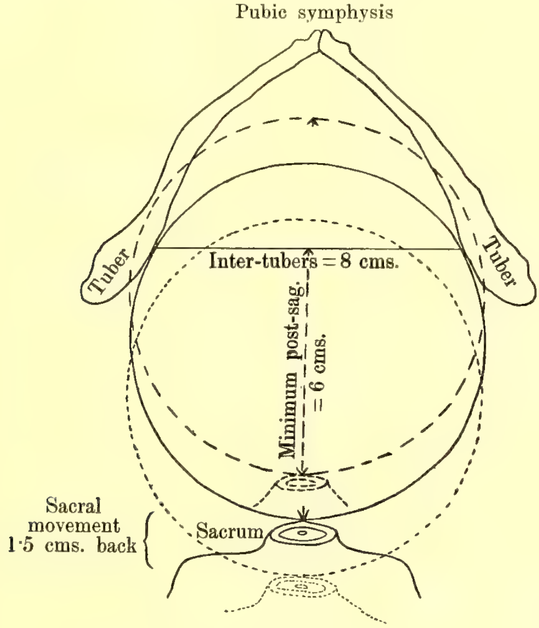


DIAGRAM III. Outlet.

Circles represent the average fetal head circumference, 9·5 cms.
Inter-tuberal diameter of 8 cms. requires a posterior-sagittal, 7·3 cms. or more ———— .
Normal sacral movements lengthen the posterior-sagittal, 1·5 cms. - - - - - - .
Reduction of the posterior-sagittal to minimum, 6 cms. – – – – – – .

The antero-posterior diameter of the outlet (Diagram I, P—S) was measured from the nearest point of the inner surface of the pubic symphysis to the tip of the sacrum. The normal rotation of the sacrum on its axis, usually the second sacral segment, may during parturition lengthen this diameter, as well as the posterior sagittal, a variable amount, from 1·5 to 2 cms. (9). Such lengthening would allow the head to pass the tuberosities more posteriorly; this in turn would bring the points of resistance on the tuberosities a variable distance back on those curved bony prominences, and, depending on this curve, the points of impingement of the head would fall a variable distance further out. Consequently we see that the moving backward of the tip of the sacrum for this circular passenger enlarges the available space not merely directly in proportion to the distance backward, but more nearly by the square of that distance. The figures of the measurements taken do not include such increase in the outlet space from the mobility of the sacrum (Diagrams I and II).

The following table gives the figures of the outlet in convenient condensed form :—

Arch (pubic)	Antero-posterior	Freq.	Inter-tuberal	Freq.	Posterior-sagittal	Freq.
Roomy ... 71	9– 9·9 cms.	3	8– 8·9 cms.	40	6– 6·9 cms.	42
Medium-roomy 4	10–10·9 „	48	9– 9·9 „	73	7– 7·9 „	102
Medium ... 92	11–11·9 „	89	10–10·9 „	85	8– 8·9 „	62
Medium-narrow 5	12–12·9 „	58	11–11·9 „	16	9– 9·9 „	9
Narrow ... 44	13–14·66 „	19	12–12·75 „	3	10–11·7 „	2
Average	11·59 cms.		9·79 cms.		7·56 cms.	
Maximum	14·66 „		12·75 „		11·7 „	
Minimum	9 „		8 „		6 „	

Small outlets, specimens in which the "fetal head" scale measured 9·5 to 9 cms. in diameter, numbered twenty. As may be seen in the diagrams, the backward movement of the tip of the sacrum to the normal amount would enlarge these outlets sufficiently to allow the easy passage of the average fetal head*, and can thus be considered efficient pelvic outlets.

The ischial spines, commonly the point of narrowest diameter passed by the fetal head in its descent through a normal pelvis, may well be an important factor in the mechanism of labour. In my series the attempt was at first made to measure this diameter, but in so many specimens it was found that this prominence was broken off in part or entirely, that anything like accuracy was impossible. It was decided, therefore, that any figures so obtained would be deceiving rather than helpful, and none are given.

An index of the pelvic outlet was estimated as follows :—

$$\textit{Outlet Index} = \frac{\text{inter-tuberal diameter} \times 100}{\text{antero-posterior of outlet}},$$

average 84·26, maximum 108·9, minimum 64.

The two diameters used in this index, however, lie in different planes, and can, therefore, hardly be considered a true measure of this irregular exit space (cf. Plate VI, 7 and Diagram I, *S—P* and *S—T*). The distance between these two diverging planes was estimated at the point opposite the tuberosities by erecting a perpendicular (*O—T*) with the antero-posterior diameter (*P—S*), as a base line, to the inter-tuberosities line (*T*). The distance from the pubes along the base line to this perpendicular was also measured to show how far back the tuberosities were placed (Diagram I, *P—O*). The last of these distances (*P—O*) shows approximately how far posterior to the pubes the broadest part of the fetal head may be forced to pass, when the tuberosities are sufficiently narrowed. The first

* Biparietal diameter 9¼ cms., sub-occipito-bregmatic 9½ cms., Farabeuf and Vernier.

measure $(O—T)$ represents the amount to which the birth canal is prolonged downward and backward, when the tuberosities are thus narrowed.

In *recapitulating* our findings in respect to the outlet it is fair to say, first that this is a complicated space, the points of obstruction being on different and diverging planes, varying in divergency, thus multiplying the number of elements in the problem. Some of these factors I have tried to elucidate, but others, including the ischial spines and the variations in size, shape, and malleability of the fetal head, must be left for further studies on the living, and for nature's final test of labour.

By means of the "fetal head" scale the points of impingement on the tuberosities have been determined with comparative accuracy on this series of dry bones. These points on the tuberosities were found usually to be on the inner lip of the tuberosity at a varying distance from the symphysis. This variation was considerable and depended on three factors, the curve of these bones, their distance from each other, and their distance from the sacrum. By this means it was also found that some pelves which appeared to have small generally contracted outlets were in reality quite passable.

The ischial spines, though not available here, must be considered as probably an important factor in the mechanism of labour at or near the outlet.

Finally, emphasis is laid on the importance of the space between the tuberosities, and, when this is reduced, on the available space behind, measured by the posterior-sagittal diameter. This combination is probably the most accurate practical measure to show the significance of the variations of the outlet.

The *separate bones* of the pelvis were measured. The length of the *innominate bone* was obtained by means of the graduated measuring board and block, thus giving the maximum distance from crest to tuberosity. The width of the ilium was measured with the "compas glissière" from the anterior to the posterior superior spine. An *index* for the larger innominate was figured $= \frac{\text{breadth} \times 100}{\text{height}}$. Also a

$$\textit{Pelvic index} = \frac{\text{height (highest innominate)} \times 100}{\text{breadth (inter-crests)}}.$$

The *sacrum* owing to its position of importance and to its great variability is a most interesting bone. Its shape often modifies markedly the pelvic cavity. Its height was taken from the middle of the anterior surface of the promontory to the anterior surface of the tip of the sacrum. The maximum breadth was taken with the instrument parallel to the anterior surface of the bone.

$$\textit{Sacral index} = \frac{\text{breadth} \times 100}{\text{height}}.$$

Four "observations" were added. 1. The number of sacral segments. 2. The sacral curvature was estimated as "slight," "moderate," or "pronounced." 3. The segment noted at which the curve began. 4. False promontories were noted and used for the measure where the cavity was involved in the measurement.

Summary. The features brought out by this investigation are: 1. The variation in the size, shape, and type of the female American Indian pelvis as a whole. 2. The variation in the pelvic inlet and 3. outlet. 4. The variation in the number of sacral vertebrae. 5. The frequency of a false promontory.

1. The variation in *size* of the normal female pelvis is considerable, as is illustrated in Plate I, 1.

The *shape* of the pelvic cavity also varies much, as is suggested by Plate II, 2, comparing a flat inlet with a rounded one; and Plate VII, 9 contrasting a wide outlet with a narrow one.

The "male" type or high, narrow-arched pelvis is shown in Plate III, 3, compared with the "female" or extremely shallow, broad-angled pelvis. The "male" type suggests a birth canal with a small bore, long cylindrical cavity ending in a narrow-arched small outlet. The effect on the mechanism of labour is to increase its difficulty proportionately.

The determination of the sex of a pelvis is not always easy, and in a small percentage (perhaps 1 to 3 °/₀) of cases is next to impossible, even with the aid of the whole skeleton. The breadth of the great sacro-sciatic notch was found to be the most reliable guide. The acuteness of the sub-pubic angle was next in usefulness. The thickness of the bones and finally the skull and long bones were referred to in doubtful cases (10). No specimen was included unless determined by these signs to be female.

2. Considerable variation is seen in the diameters of the *inlet.* Michaelis (11) reported 1,000 cases carefully studied as to pelvimetry and the results of labour. Litzmann (12) continued the work, reporting a second 1,000 cases. The standard determined by these men has been accepted throughout the world. Litzmann (12) considers all pelves contracted if the conjugata vera is 10 cms. or less in a generally contracted pelvis, and 9·5 cms. or less in a flat pelvis. By this standard my series of Indian pelves show 63 "contracted pelves," or 29 °/₀, as follows:—

Generally contracted 9 to 10 cms.	6	Flat 9 to 9·5 cms.	43
" " 8 to 8·9 "	1	" 8 to 8·9 "	11
— — — —	—	" 7·5 to 7·9 "	2

A pelvis was classed "generally contracted" if the transverse diameter was less than 12 cms., at least one centimetre below the average: with, at the same time, an obstetric diameter of 10 cms. or less. A pelvis was considered "flat" if the transverse diameter measured 12 cms. or more with the obstetric diameter 9·5 cms. or less.

Three only of the seven generally contracted pelves showed the inter-cristal and inter-spinous diameters reduced more than 1·5 cms. below the average, and thus might be said to "suggest" lateral contraction within normal limits. The other four showed no such suggestion of contraction.

The flat pelves, 56 in all, show that at least slight contraction of the inlet is to be expected in one pelvis in every four, while moderate to serious narrowing of this all-important space is to be expected about once in twenty pelves.

3. The *outlet* showed wide variation in size and shape. By the use of the "fetal head" measure it was determined that no pelvis, in all probability, was too small to pass the average fetal head. It was found that a narrowing of the inter-tuberosities diameter was the most important single factor in reducing the size of the available space. Next in importance was the shortening of the space behind this line, or in other words a reduction of the posterior-sagittal diameter, found with but slightly greater frequency in pelves in which the sacrum contained an increased number of segments. The combination of these two factors, short inter-tuberosities and posterior-sagittal diameters, was essential to reduce seriously the efficiency of the outlet. The normal movement of the sacrum, allowing the tip to swing backward, enlarged the available space considerably. The diameter of the ischial spines, though probably of great importance in the efficiency of a pelvis, was not available in this series of ancient pelves.

4. A *numerical variation* of the *sacral vertebrae* was noted in 47 pelves or 21·7 %, or one in every five pelves. The number of segments ranged from four to six (see Plate IV). By an increased number of segments is meant sacra in which the six segments were all sacral in character, or those in which there was a transitional vertebra, whether lumbar or coccygeal in character. Such a classification, as was pointed out to me by the late Professor T. Dwight, is more practically useful than anatomically correct. Pelves with small outlets were found slightly more often among those with an increased number, and it is possible that this increase in segments may be one small factor in reducing the size of the outlet. Aside from the possibility of slightly infringing on the outlet space, in rare cases the numerical variation appears to play no important part in the variation in size and shape of the pelvic cavity (Plate VII, 9 and 10).

5. *False* and *double promontories* were found in twenty pelves, about one in every eleven pelves. These false promontories varied from a marked prominence of the second sacral vertebra (Plate VI, 6), an equal prominence of the first and second sacral segments, to a projection of the top of the last lumbar vertebra beyond the sacrum nearer to the pubes (Plate V, 5). In nearly all cases false promontories were associated with transitional vertebrae or an increased number of sacral segments. The apex of the lumbo-sacral bend falls at a point proportionally distant, in all probability, from the sacro-iliac attachment or the "vertebra fulcralis" of Welcker (13), and in these transitional cases this distance brings the point on another vertebra than the usual one. False promontories occurred in four specimens classed as moderately generally contracted and in one with a small outlet, but in no other "contracted" pelvis. That is a little more often than the general average, but in all probability has no special significance. From the general appearances as well as from the measurements it seems fair to say that the false promontory has no appreciable effect on the pelvic cavity.

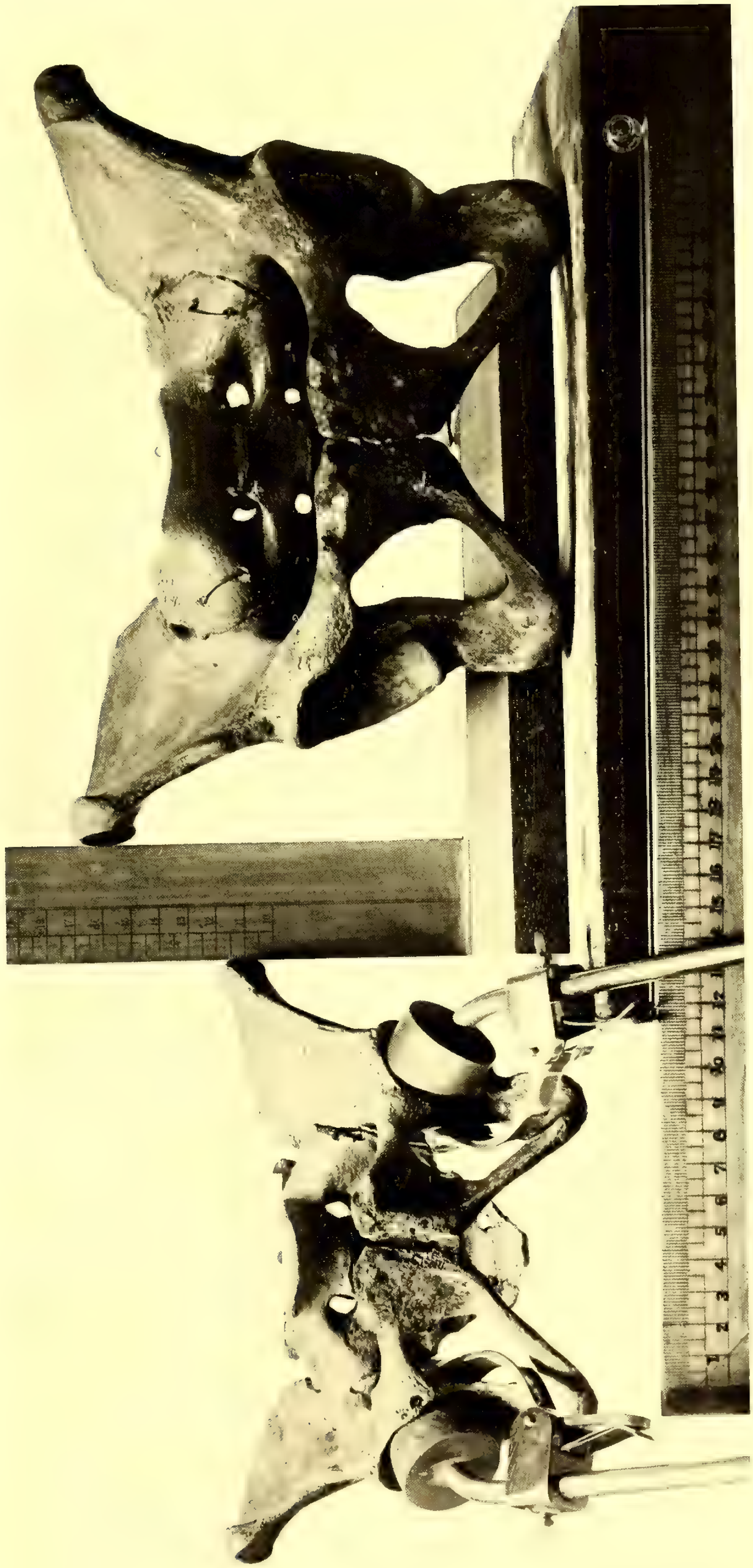
Congenital dislocation of the *hip* was found in one specimen (#10). The displacement, of the right femoral head only, was upward and backward where a new socket was formed on the ilium. The deformity produced a tilt to the pelvis of approximately 10 degrees downward to the right. A new acetabulum was formed, and the femoral head and neck showed practically no change. As will be seen by the measurements no alteration is found in the pelvic cavity in spite of this deformity.

In conclusion I would emphasize that a "normal" standard for the pelvis should include not only an average, but also a series of measures graded from minimum to maximum. A large series would seem necessary in order to include the wide variations found among non-diseased pelves. Owing to these great differences found among the specimens of such a comprehensive series, less stress, I think, should in future be put by the anthropologist on those small differences seen in a few or even in a small series of pelves of separate tribes or races, but rather the close resemblances should be emphasized in contrast to the great differences noted between animals and man. And the changes might be further studied and traced which have followed the assumption of the erect attitude.

Such changes as the differences in the sexes are receiving the attention of the embryologist (14). The anatomist is studying the variations of the spine and their significance. The obstetrician is constantly seeking more light on the interpretation of peculiarities of this complex bony structure. It was my desire, therefore, to suggest to the anthropologist an application which may be made of his work of collecting and classifying, identifying and interpreting such valuable material as is here used, and also to direct his attention to the applied anatomy and physiology of the pelvis: that is the function of weight-bearing and child-bearing.

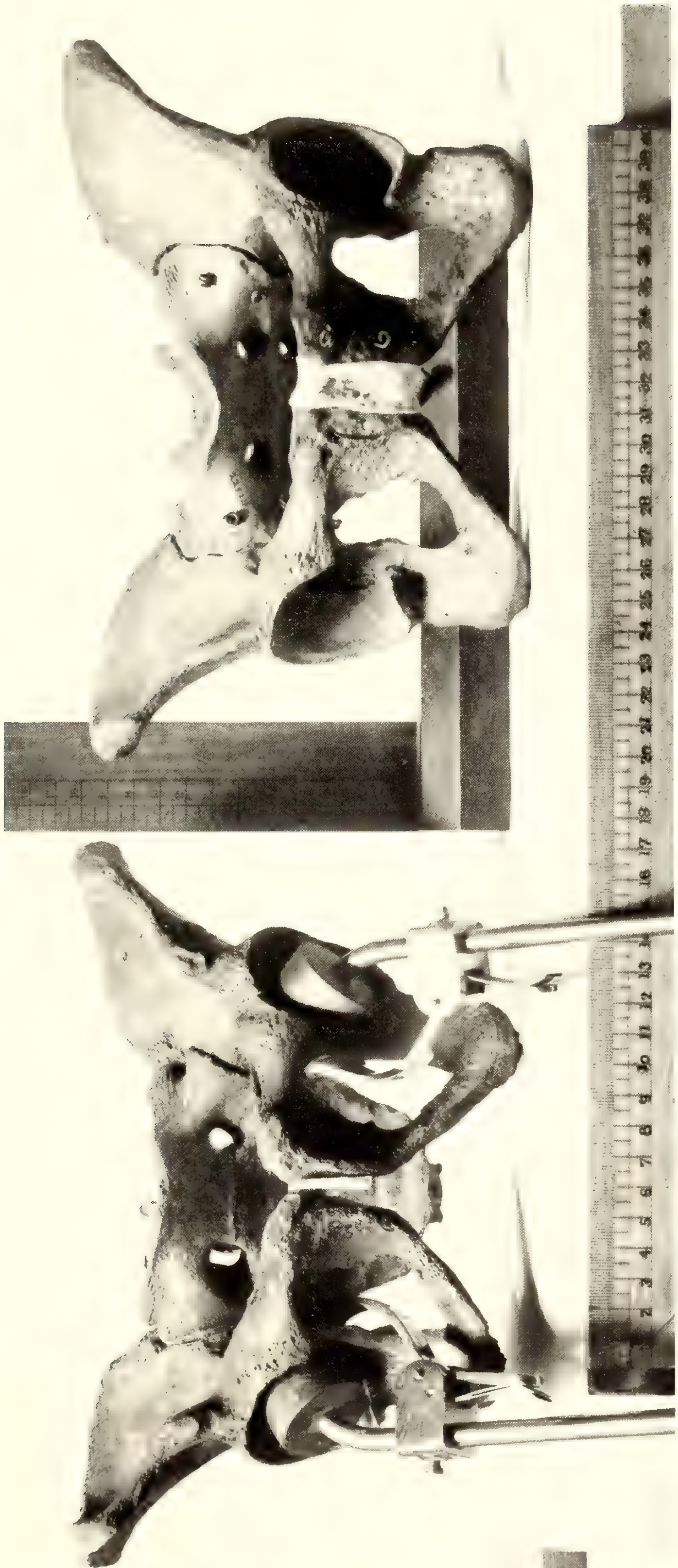
The popular idea has been that the function of child-bearing among the American Indians was always efficient and easy. Engleman (15) confirms this idea by saying that labour as a rule among North American Indians is short and easy, averaging two hours. As civilization is approached, however, labour becomes more extended. Thus half-breeds, as the modern Mexican Indian, average three to four hours. He further states that accidents during labour are rare when women do not marry out of their tribe, for the child's head is in proportion to the pelvis. But deviation from the natural state, he continues, brings difficulty. The example is given of the Umpqua tribe, who have intermarried with whites and have died, it is stated, during labour in consequence of the disproportion between the larger head and the ordinary pelvis. When the father of the child, however, was also an Umpqua no such trouble is known to have occurred.

My series of dried specimens bears out in general this clinical evidence, for most labours might well have been short and easy so far as the pelvis was concerned. About one-fourth, however, would require a rather smaller fetal

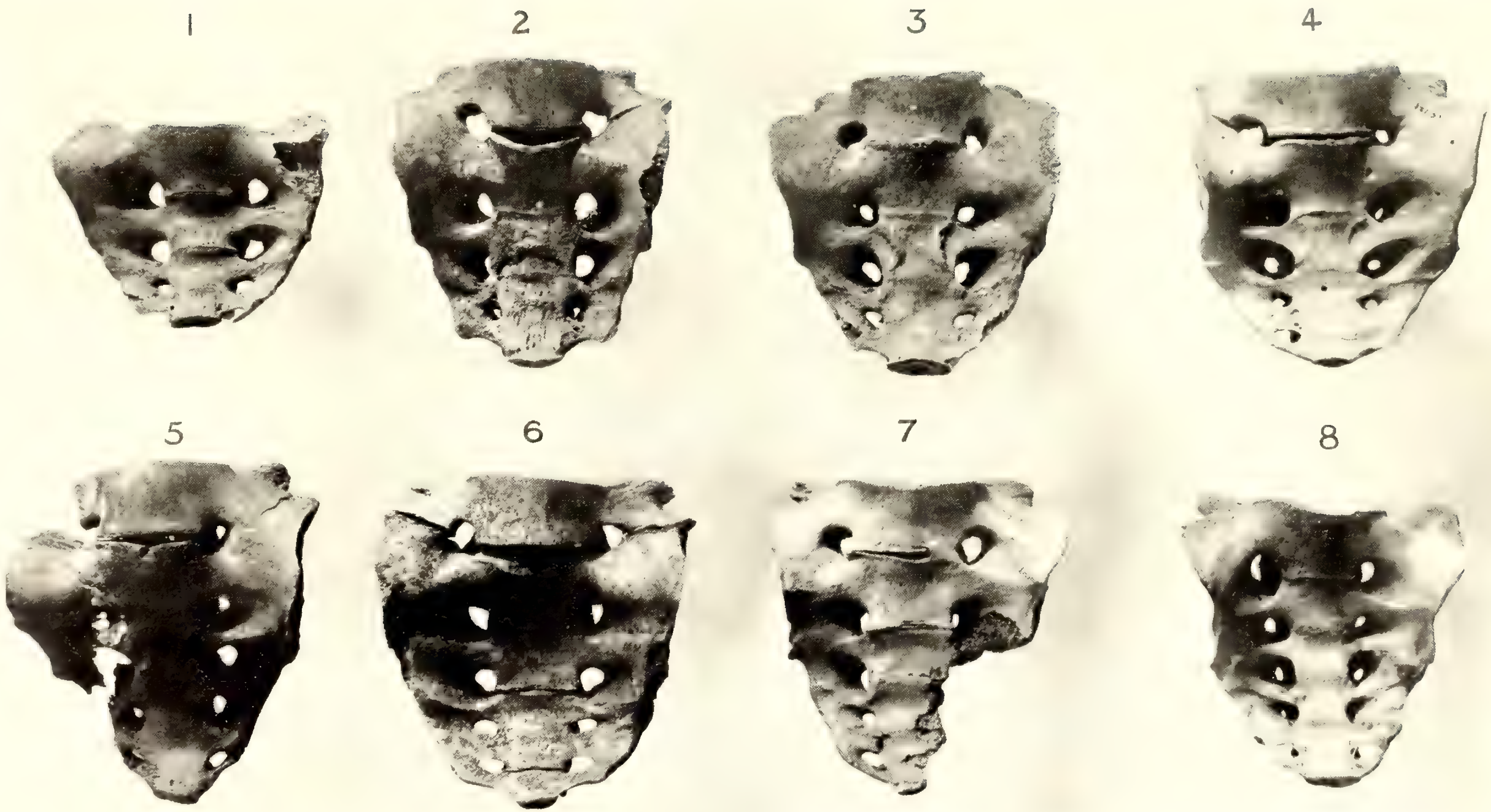


1. Variation in Size.—Small, # 62, Arizona.—Large, # 107, Sioux.

2. Variation in Shape.—Rounded, # 86, Kentucky.—Flat, # 73, Comanche.

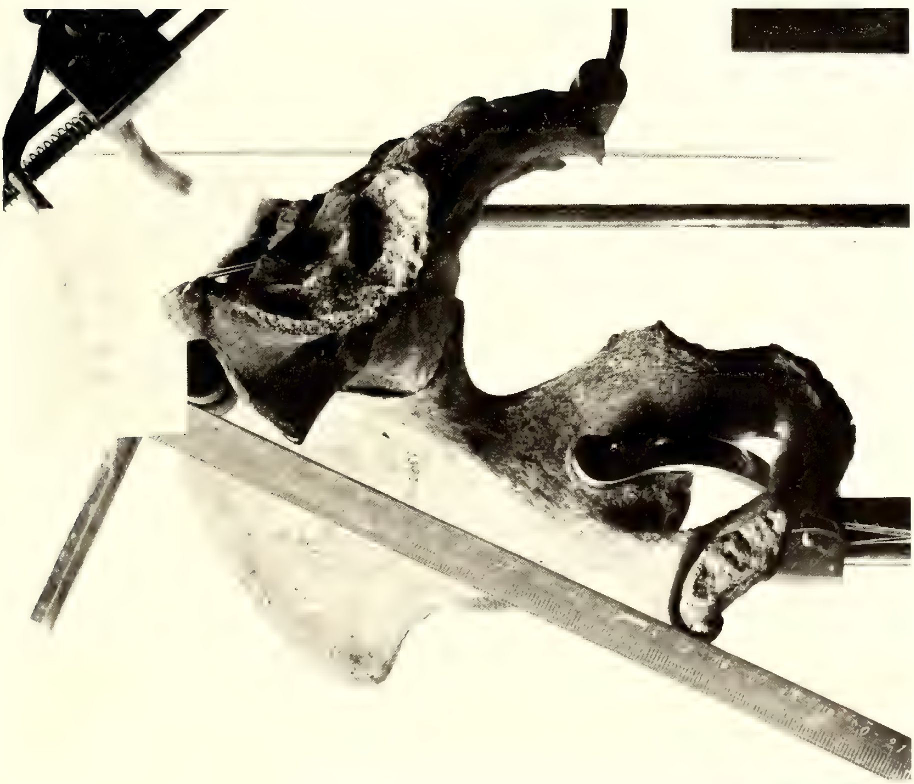


3. Variation in Type.—"Male," #46, Apache, squaw.—"Female," #95, Pah Ute.

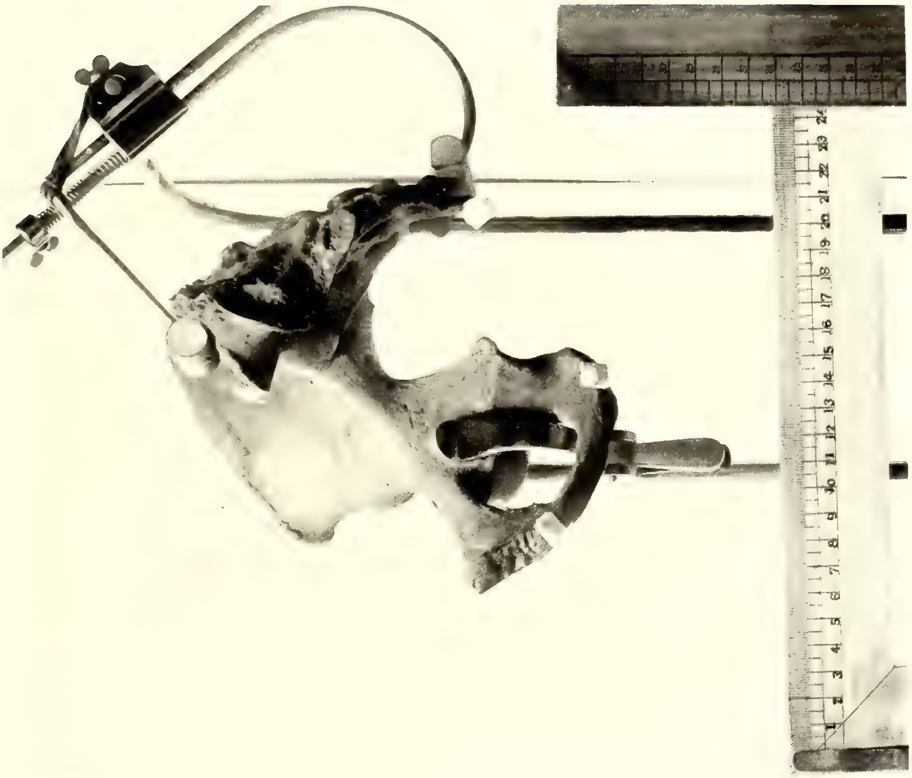


4. Numerical Variations of Sacral Segments, Progressive Series. 1. ♯ 60, 4 segments. 2. ♯ 51, $4\frac{1}{2}$ segments. 3. ♯ 57, 5 segments. 4. ♯ 109, $5\frac{1}{2}$ segments. 5. ♯ 106, $5\frac{1}{2}$ segments. 6. ♯ 94, 6 segments. 7. ♯ 74, 6 segments. 8. ♯ 99, 6 segments.

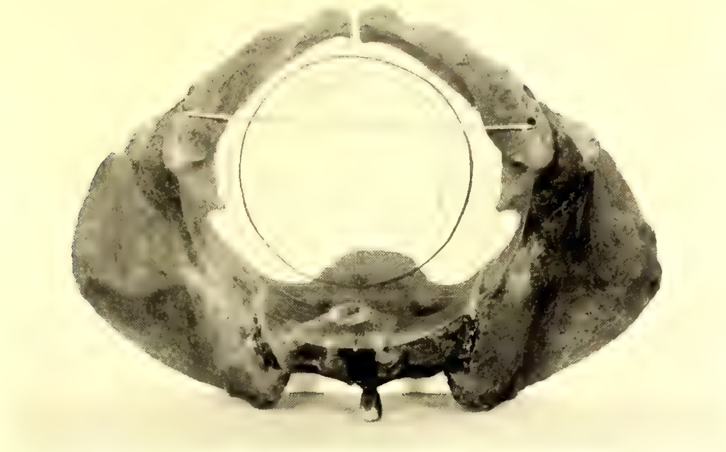
5. "False Promontory" made by first lumbar vertebra. ‡77. Apparatus for holding pelves (Author).



6. "False Promontory" made by the second sacral vertebra. # 100.

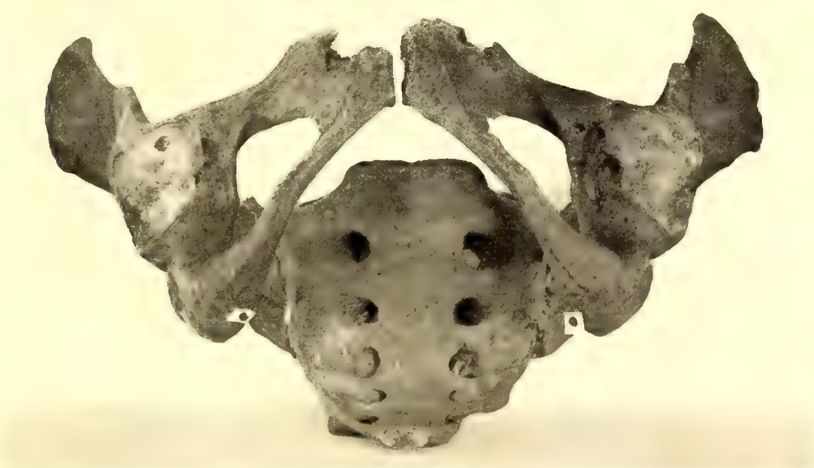


7. # 106, "Outlet" small, marks show points of measurement and planes of pelvis.



8. Celluloid circular scale for Outlet, dark circle diameter 9 cms., string marks tuberosities. Pelvis ♯15 California, 13,232, six sacral vertebrae.

9. Variation of Outlet. String marks intertuberal diameter. Wide: ♯13, California 13,554, Intertuberal 12·75, Antero-posterior 14·66 cms. Narrow: ♯34, Massachusetts, 47,998, Intertuberal 8, Antero-posterior 12 cms.



10. Six sacral vertebrae, all sacral in character. Roomy Outlet. Tuberosities marked. 10·1 cms. Antero-posterior 1·75 cms. California, 13,232 (same as photo. 8).

head than the average, as found in civilized countries and peoples, in order for labour to be short and easy. About one in ten would require a small and perhaps malleable head to make successful labour possible. Observations on the heads of new-born pure-blooded Indian infants as to size and malleability would be a welcome supplement to our present knowledge of pelvic efficiency.

BIBLIOGRAPHY.

(1) HRDLICKA, A. U.S. National Museum, personal communication.

WILSON, T. The antiquity of the Red Race in America. *Reports, Smithsonian Institute*, 1897, p. 1041.

(2) TOPINARD, P. Éléments d'Anthropologie générale. Paris, 1885.

——. Des proportions générales du bassin chez l'homme, &c. *Bull. de la Soc. d'Anthropologie*, 1875, pp. 504—521.

(3) VERNEAU. Le bassin dans les sexes et dans les races. Paris, 1875.

(4) TURNER, W. Reports H.M.S. *Challenger*, Vol. XVI, Zoology, Part II.

(5) BAUDELOQUE. L'art des accouchements. Nouvelle éd., 1789, T. I, pp. 76—90.

(6) SCHEFFER. *Monatsschr. f. Geburtsk.*, 1868, Bd. XXXI, pp. 299—309.

(7) WILLIAMS, J. W. Obstetrics, 1908, p. 688.

(8) ——. Frequency, Etiology, and practical significance of contractions of the pelvic outlet. *Surgery, Gyn. and Obst.* Vol. VIII, pp. 619—638. June, 1909.

(9) KUETTNER, O. V. Hegar's Beitraege, 1898, Bd. I, pp. 211—228.

(10) DWIGHT, T. The size of the articular surfaces of the long bones as characteristic of sex. *Am. J. Anat.* Vol. IV, pp. 19—31.

(11) MICHAELIS. Das enge Becken. Leipzig, 1851.

(12) LITZMANN. Die Geburt bei engen Becken. Leipzig, 1884.

——. Die Formen des Beckens. Leipzig, 1861.

(13) WELCKER. Zur Lehre von Bau u. Entwicklung der Wirbelsäule. *Zoolog. Anz.*, 1878, Nr. 13.

(14) THOMSON, A. The Sexual Differences of the Fetal Pelvis. *Jour. Anat. and Physiol.* London, 1899, Vol. XXXIII, pp. 35—80.

(15) ENGLEMAN, GEO. J. Labour among Primitive Peoples. 3rd ed. 1884, St Louis Mo., pp. 7—9.

Table of Pelvic Measurements of

No.	Peabody Museum No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Inlet: Oblique Diameter	Inlet: Obstetric Conjugate	Inlet: Difference	Inlet: Index	Pelvic: Antero-posterior Diameter	Pelvic: Fetal Head	Pelvic: Inter-tubers	Pelvic: Posterior Sagittal	Pelvic: Pubes to Perpendicular
1	11,970	Tennessee	25·5	23·25	13·1	12·7	11	1·7	84	12·5	10·75	9	7·3	*7*
2	57,506	Ohio	27·25	25·25	13·5	12·5	11·25	1·25	81	10·8	10·5	9	8·3	4·5
3	26,630	,,	26	23·75	12·5	13·25	11·75	1·5	94	11	9·5	8·1	6·8	5·3
4	26,989	,,	26·5	23·75	12·75	11·3	9·5	1·8	74·5	11·6	12	10·3	8·3	4·5
5	13,551	California	24	20	13	12·2	11·2	1	86·2	10·6	11	9·2	7·7	4
6	13,238	,,	23·75	20	11·75	11·75	10·75	1	91·4	11·15	11	10	8	4·2
7	13,240	,,	27	24·5	14	11·4	10·1	1·3	72	11	11	10·1	7·7	4·2
8	13,239	,,	26·25	22·75	13·25	12·3	10·6	1·7	80	11	9·5	9	6·7	5·8
9	13,286	,,	25·5	20·7	13·25	12·3	11	1·3	83	10·8	10	8·85	7·4	4·5
10	13,448	,,	25	22	13	12·1	10·4	1·7	80	10·9	11	9·9	7·6	4·5
11	57,883	,,	27·5	25	13·25	11·5	9·5	2	71·1	13·75	12·5	10	*10·5*	4·5
12	13,553	,,	25·5	20·5	13·75	11·5	10·2	1·3	74·2	13	10	9·1	7·5	4·85
13	13,554	,,	26·25	21	14·25	11·5	9·5[a]	2	66·6	14·66	13	12·75	8·24	3·6
14	12,248	,,	24	22	12	12·9	10·9	2	90·8	13	12	10·35	8	6·2
15	13,232	,,	26·75	22·5	13·5	11·5	9·35	2·15	69·3	11·75	10·5	10·1	7·1	6·15
16	13,545	,,	24·75	22	12·75	12·5	10·5	2	82·3	11	10·25	9·9	7	5
17	13,546	,,	27[a]	24·5[a]	13	14·7	13	1·7	100	12·75	11	10·2	7·5	5·6
18	13,547	,,	25·5	23[a]	13	12·3	11[a]	1·3	84·6	11·25	11	10·1	7·8	4·6
19	13,548	,,	26·25	23	13·5	12·2	11	1·2	81·4	12	10	9	7·75	6
20	48,009	Mexico	24·75	21	13·9	11	9·8	1·2	70·5	12·25	10·5	9·5	7·3	6·25
21	57,782 ?	California	27	24·25	14	12·2	10	2·2	71·4	10·75	12	11·7	7·2	5·6
22	58,143	Iroquois, N. Y.	27	25·5	13·25	12·4	10·2	2·2	73·6	*9*	10	9·8	7·5	4·65
23	8,359 ?	Flat Head	27·25	25·5	14·25	12	10·25	1·75	71·9	11·1	11	10·3	6·6	5·5
24	57,458	New York	25·25	23·5	12·5	11·7[a]	9·5[a]	2·2[a]	76	12	11	9·1	8·1	5·1
25	2,347	Kentucky	25·25	22	13	12	10·5	1·5	80·8	12	11·5	10·9	8·2	5·1
26	2,346	,,	28·5	26·25	14	13·2	11·5	1·7	82·1	13	10·5	9·5	7·7	6·5
27	2,350, III	,,	27·5	23	13·5	12·25	10·75	1·5	79·6	11·7	10·5	8·8	8·2	4·6
28	11,972	Tennessee	23·5	22	13	13·5	12	1·5	92·3	11·75	11·5	10·2	8·6	4
29	11,861	,,	27[a]	24[a]	14	12·6	11	1·6	78·6	11	10·75	10·1	7·8	4
30	57,512	,,	25·5	23·25	14	11·5	9·8	1·7	70	12·25	12	10·3[a]	8·8	4·6
31	27,213	,,	26	24·25	13	11·5	10·25	1·25	78·8	10·25	10	9·7	7	4·15
32	32,435	Massachusetts	26	23·25[a]	13	11·5	10	1·5	77	11·2	10	8·8	6·5[a]	5·7
33	10,262	,,	23	20	12	12·5	11	1·5	83·3	11[a]	*9*	*8·1*	*6*	6·4[a]
34	47,998	,,	28·5	*27*	13·5	13·1	11	2·1	80·4	*12*	10	*8*	*8*	6·1
35	57,779	Arizona	25	23·3	13	9·8	*7·75*	2·05	59·6	11	10·5	10	6·5	5·4
36	58,005	Ohio	25	22·5	12	12	11	1	83·3	12	12·5	10·8	9·5	4
37	58,022	,,	27·5	25[a]	12	*14·8*	*14*[a]	0·8	*107·7*	13	12	10·3	8	5·5
38	58,049	,,	26	24	13	11·7	*9*	2·7	69·2	12·5	11	10·4[a]	6·5[a]	7[a]
39	58,056	,,	27[a]	24·25[a]	12·75	13·2	11·5	1·7	90·2	11·25	11	9·9	8	4·7
40	58,023	,,	26·25	23	13	14·1	13	1·1	100	12	12	12	7·8[a]	6
41	58,057	,,	23[a]	20·75[a]	12	13·2	10·25	2·95	85·4	10·25	9·75	8·8	7·6	4·5
42	58,453	,,	25·75	23·5	13·5	11·2	10	1·2	74	12	11·5	10·2	8	5·5
43	58,463	,,	28	25	14·5	12·3	10·5	1·8	72·4	12·3	12	10·6	9	5
	U. S. N. M.													
44	227,463	Alaska	23·1	20·5	11·2	12	9·7	2·3	86·6	10·3	*9*	10·6	*6*	5·4
45	225,473	Apache	25·75	23·8	13·9	12·7	10·2	2·5	75·5	12·2	11·5	10·7	8·7	6
46	228,361	,,	25·5	22	11·5	11·9	10·2	1·7	88·7	10·2	9·5	8·5[a]	7·2	3·9
47	226,290	Arizona	26·5	23	13·8	11·8	9·5	2·3	69·5	12·2	9·75	8·4	7·1	6·3

1. One sacral segment probably from coccygeal end. 3. Last lumbar articulates with left sacral ala. 5. Four sacral segments. No evidence of more at either end. 6. First sacral transitional. False promontory. 7. False promontory. 10. Congenital dislocation of right hip. Two sockets. Inlet diagonals, right 12·3, left 12·3. 11. Coccyx ossified to sacrum (not included), distorted. 12. False promontory. 13. Wide outlet (Plate VII, 9). 15. Well proportioned symmetrical sacrum (Plate VII, 8 and 10). 17. Coccyx ossified to sacrum (not included). 20. Tapering sacrum.

(Note: "a" means an approximate measurement often due to a broken bone. Italics signify high or low value or some peculiarity referred to in text.)

American Indian Squaws.

Outlet				Innominate Bone						Pelvic Index	Sacrum					
Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index		Height	Breadth	Number Segments	Curve	Segment begins at	Index
3·3	3	Medium	80	19·8	14	70·8	19·8	14	—	77·6	10	11·5	*6*	Moderate	1st	115
4	3·3	M	83·3	20·6	15·2	—	21·1	15·6	73·9	77·4	11·1	12·4	5	M	3rd	112·6
3·3	2·9	Narrow	73·6	19·2	13·8	—	19·4	13·7	70·6	74·6	11·6	11·1	5	Slight	1st	95·7
3·5	3·4	Roomy	88·6	19·9	15·4	77·4	19·9	15·4	—	75·1	9·9	11·4	5	M	1	115·1
3·7	2·3	M	86·8	18·5	14	—	18·5	14·3	77·3	78·7	*6·3*	11	*4*	Pronounced	1	*174·6*
2·6	4·3	R	89·6	18·3	14·5	—	18·6	14·4	77·4	78·3	10·5	10·5	5	M	2	100
3	2·4	R	91·8	19	14·75	—	19·2	14·65	77·6	71·1	10·95	12·6	5	S	4	115·1
3	3·5	N	81·8	20·4	15·5[a]	76	20·4[a]	15·5[a]	—	73·9	10·1	12 25	5	S	3	121·2
3	3·4	N	81·8	18·8	14[a]	—	18·8	15	79·8	73·7	9·7	12·1	5	S	1	124·7
3·6	2·65	R	90·8	19·7	13·5	68·5	19·4	13·8	—	74·4	9·3	11·9	5	S	1	127·9
4·5	1·7	M	75	19·9	16	80·4	19·9	15·6	—	72·4	8·6	10·95	5	P	1	127·3
2·75	2·2[a]	M	82·5	19	14·4	75·8	19	13·75	—	74·1	10·25	11·1	5	S	4	108·3
3	?	R	87	18	14·5[a]	—	18·6	14·5[a]	72·6	70·9	9·1	12·2	5	M	2	134·1
3·75	3	M	79·6	19·1	14·8[a]	—	19·8	14·5	73·2	82·5	9·8	11·9	5	M	3	121·4
3·25	2·8	M-R	85·9	19·1	14·6	76·5	19·1	14·4	—	73·6	10·6	12	*6*	M	1	113·2
3	3·25	R	90	19	15·5[a]	76·3	19	14·5	—	76·8	9·65	11·75	5	M	1	121·7
3	2·75	M	80	20	15·2	—	20·5	15·2[a]	74·1	75·9	11·3	12·2	5	M	1	107·9
3	2·5	M	89·7	18·6	14·4	77·4	18·6[a]	14·4[a]	—	72·9	8·1	12·3	5	M	1	151·8
3·5	2·7	M	75	18·4	14·6	—	18·7	14·4	77	71·2	8·4	12·3	5	M	1	146·4
3·25	2·35	M	77·5	19·2	14·8	77·1	19·2	14·6	—	77·6	9·1	11·1	5	P	1	121·9
4	2·7	R	108·8	19 8	14·4	72·7	19·7	14·5	—	73·3	9·6	12·6	5	M	1	131·2
3·25	3[a]	M	*108·9*	19·5	14·2[a]	72·8	19·5[a]	14·2	—	72·2	10·6	12·5	5	S	2	117·9
3	3	M	91·1	19	15	—	19·2	15	78·1	70·5	8·9	12·1	5	M	1	135·9
3·5	3·6[a]	M	75·8	18·6	13·7	—	19·1	14·1	73·8	75·6	10·5	10·85	5	M	1	103·3
3·5	2	R	90·8	17·8	14	78·7	17·8	14	—	70·5	9·9	10·8	5	S	3	109·1
3	3·1	M	73·1	20·8	15·5	—	21·2	15·5	72·1	74·4	11·6	12·5	5	S	2	107·7
3·75	2·9	M	75·2	19·6	15·3	78·1	19·6	15·2	—	71·3	8·4	11·9	5	M	1	141·6
2·6	2·5	M	86·8	19·1	14	—	19·7	13·7	69·5	83·8	9·9	11·2	5	M	3	113·1
2·5	3·3	R	91·8	21·1	15·1	—	21·4	15·1[a]	70·5	79·3	11·25	12·7	5	M	2	112·9
3	2·8	R	84·1	19·4	15·2	78·8	19·4	15·2[a]	—	76·1	12·9	11·1	*6*	S	5	*86*
2·5	3·5	M-R	94·5	18·4	14·5	73·3	18·4	14·5	—	70·8	*13·2*[a]	11·4	*6*	S	3	86·3
3·5	2·9	N-M	78·6	19·2	13·8[a]	71·3	19	13·8	—	73·8	10[a]	11·7	5	M	1	117
2·5	3[a]	N	73·6	18	12	—	18·2	13·8	75·8	79·1	10·6	10·3	5	S	1	97·1
4·35	3·4	N	66·6	20·5	16	78	20·5	15·8	—	71·6	10·2	12·2	5	M	3	119·6
3·7	2·5	M	90·9	18·7	13·8	—	18·8	14·1	75	75·2	9·5	11·4	5	M	4	120
4	2·6	R	90	18·6	15	—	18·8	14·5	77·2	75·2	11·1	10·9	5	S	1	98·2
3·75	3·2	M	79·2	*22·2*	16·5	—	*22·5*	16·5[a]	73·3	81·8	11·1	12·6	5	P	1	113·5
3·5[a]	3·1	M	84·9	19	13·5[a]	—	19·2	13·5	70·3	73·8	10·1	11·2	5/6	S	3	110·9
3·5	2·8	M	88	20·2	15·2	—	20·7	15·2[a]	73·4	76·7	10·2	12·2	5	M	2	119·6
3	2·6	R	100	21	15·5	—	21·3	15·6	73·2	81·1	11·1	12	5	M	3	108·1
3·5	2·9	N	85·8	18	13[a]	72·2	18[a]	13[a]	—	78·2	8·8	11·7	5	M	4	133
3·75	2·7	M	85	19·8	15·4	77·8	19·8	15·2	—	76·9	12·1	11·3	*6*	M	4	93·3
3·5	3·3	R	86·2	21·3	16	75·1	21·2	16	—	76·1	11	13·7	5	M	3	124·5
3	2·6	N	79·6	19·5	14·5	73·8	19·5	14·5	—	84·4	10·35	10·25	5	S	2	99
4	2·7	M	87·7	21·0	15·6	—	21·2	15·75	74·3	82·3	7·7	12·2	5	P	1	158·4
3·5	3·5	N	83·3	20·1	15	74·6	20	15·1	—	78·8	9·5	10·8	5	M	3	113·7
3·7	3·0	N	68·9	19·4	13·2	68	19·1	13·7	—	73·2	10·9	11·45	5	S	4	105

21. Ossification, left ilio-sacral synchondrosis. 26. Large pelvis, heavy bones. 30. Six sac. vert., one from coccyx (?). False promontory. 31. Six sac. vert., one from lumbar. False promontory. 33. Small pelvis, coccyx ossified to sacrum. 34. Sacral angle acute (3rd vertebra) (Plate VII, 9). 35. Narrow flat inlet. 37. Large pelvis. 38. Left sacral ala rises to 3·1 cms. above body and articulates with last lumbar. 42. Six sac. vert., one from lumbar(?). False promontory. 44. Small pelvis. "Male type." 46. "Male type" (Plate III, 3). 47. "Male type"

Table of Pelvic Measurements of American Indian Squaws.

No.	Peabody Museum No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Oblique Diameter	Obstetric Conjugate	Difference	Index	Pelvic Outlet: Antero-posterior Diameter	Fetal Head	Inter-tubers	Posterior Sagittal	Pubes to Perpendicular
1	11,970	Tennessee	25·5	23·25	13·1	12·7	11	1·7	84	12·5	10·75	9	7·3	7
2	57,506	Ohio	27·25	25·25	13·5	12·5	11·25	1·25	81	10·8	10·5	9	8·3	4·5
3	26,630	"	26	23·75	12·5	13·25	11·75	1·5	94	11	9·5	8·1	6·8	5·3
4	26,989	"	26·5	23·75	12·75	11·3	9·5	1·8	74·5	11·6	12	10·3	8·3	4·5
5	13,551	California	24	20	13	12·2	11·2	1	86·2	10·6	11	9·2	7·7	4
6	13,238	"	23·75	20	11·75	11·75	10·75	1	91·4	11·15	11	10	8	4·2
7	13,240	"	27	24·5	14	11·4	10·1	1·3	72	11	11	10·1	7·7	4·2
8	13,239	"	26·25	22·75	13·25	12·3	10·6	1·7	80	11	9·5	9	6·7	5·8
9	13,286	"	25·5	20·7	13·25	12·3	11	1·3	83	10·8	10	8·85	7·4	4·5
10	13,448	"	25	22	13	12·1	10·4	1·7	80	10·9	11	9·9	7·6	4·5
11	57,883	"	27·5	25	13·25	11·5	9·5	2	71·1	13·75	12·5	10	*10·5*	4·5
12	13,553	"	25·5	20·5	13·75	11·5	10·2	1·3	74·2	13	10	9·1	7·5	4·85
13	13,554	"	26·25	21	14·25	11·5	9·5[a]	2	66·6	14·66	13	12·75	8·24	3·6
14	12,248	"	24	22	12	12·9	10·9	2	90·8	13	12	10·35	8	6·2
15	13,232	"	26·75	22·5	13·5	11·5	9·35	2·15	69·3	11·75	10·5	10·1	7·1	6·15
16	13,545	"	24·75	22	12·75	12·5	10·5	2	82·3	11	10·25	9·9	7	5
17	13,546	"	27[a]	24·5[a]	13	14·7	13	1·7	100	12·75	11	10·2	7·5	5·6
18	13,547	"	25·5	23[a]	13	12·3	11[a]	1·3	84·6	11·25	11	10·1	7·8	4·6
19	13,548	"	26·25	23	13·5	12·2	11	1·2	81·4	12	10	9	7·75	6
20	48,009	Mexico	24·75	21	13·9	11	9·8	1·2	70·5	12·25	10·5	9·5	7·3	6·25
21	57,782?	California	27	24·25	14	12·2	10	2·2	71·4	10·75	12	11·7	7·2	5·6
22	58,143	Iroquois, N. Y.	27	25·5	13·25	12·4	10·2	2·2	73·6	*9*	10	9·8	7·5	4·65
23	8,359?	Flat Head	27·25	25·5	14·25	12	10·25	1·75	71·9	11·1	11	10·3	6·6	5·5
24	57,458	New York	25·25	23·5	12·5	11·7[a]	9·5[a]	2·2[a]	76	12	11	9·1	8·1	5·1
25	2,347	Kentucky	25·25	22	13	12	10·5	1·5	80·8	12	11·5	10·9	8·2	5·1
26	2,346	"	28·5	26·25	14	13·2	11·5	1·7	82·1	13	10·5	9·5	7·7	6·5
27	2,350, III	"	27·5	23	13·5	12·25	10·75	1·5	79·6	11·7	10·5	8·8	8·2	4·6
28	11,972	Tennessee	23·5	22	13	13·5	12	1·5	92·3	11·75	11·5	10·2	8·6	4
29	11,861	"	27[a]	24[a]	14	12·6	11	1·6	78·6	11	10·75	10·1	7·8	4
30	57,512	"	25·5	23·25	14	11·5	9·8	1·7	70	12·25	12	10·3[a]	8·8	4·6
31	27,213	"	26	24·25	13	11·5	10·25	1·25	78·8	10·25	10	9·7	7	4·15
32	32,435	Massachusetts	26	23·25[a]	13	11·5	10	1·5	77	11·2	10	8·8	6·5[a]	5·7
33	10,262	"	23	20	12	12·5	11	1·5	83·3	11[a]	*9*	*8·1*	*6*	6·4[a]
34	47,998	"	28·5	*27*	13·5	13·1	11	2·1	80·4	*12*	10	*8*	*8*	6·1
35	57,779	Arizona	25	23·3	13	9·8	*7·75*	2·05	59·6	11	10·5	10	6·5	5·4
36	58,005	Ohio	25	22·5	12	12	11	1	83·3	12	12·5	10·8	9·5	4
37	58,022	"	27·5	25[a]	12	*14·8*	*14*[a]	0·8	*107·7*	13	12	10·3	8	5·5
38	58,049	"	26	24	13	11·7	*9*	2·7	69·2	12·5	11	10·4[a]	6·5[a]	7[a]
39	58,056	"	27[a]	24·25[a]	12·75	13·2	11·5	1·7	90·2	11·25	11	9·9	8	4·7
40	58,023	"	26·25	23	13	14·1	13	1·1	100	12	12	12	7·8[a]	6
41	58,057	"	23[a]	20·75[a]	12	13·2	10·25	2·95	85·4	10·25	9·75	8·8	7·6	4·5
42	58,453	"	25·75	23·5	13·5	11·2	10	1·2	74	12	11·5	10·2	8	5·5
43	58,463	"	28	25	14·5	12·3	10·5	1·8	72·4	12·3	12	10·6	9	5
	U. S. N. M.													
44	227,463	Alaska	23·1	20·5	11·2	12	9·7	2·3	86·6	10·3	*9*	10·6	*6*	5·4
45	225,473	Apache	25·75	23·8	13·9	12·7	10·2	2·5	75·5	12·2	11·5	10·7	8·7	6
46	228,361	"	25·5	22	11·5	11·9	10·2	1·7	88·7	10·2	9·5	8·5[a]	7·2	3·9
47	226,290	Arizona	26·5	23	13·8	11·8	9·5	2·3	69·5	12·2	9·75	8·4	7·1	6·3

No.	Outlet: Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Innominate Bone: Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index	Pelvic Index	Sacrum: Height	Breadth	Number Segments	Curve	Segment begins at	Index
1	3·3	3	Medium	80	19·8	14	70·8	19·8	14	—	77·6	10	11·5	*6*	Moderate	1st	115
2	4	3·3	M	83·3	20·6	15·2		21·1	15·6	73·9	77·4	11·1	12·4	5	M	3rd	112·6
3	3·3	2·9	Narrow	73·6	19·2	13·8		19·4	13·7	70·6	74·6	11·6	11·1	5	Slight	1st	95·7
4	3·5	3·4	Roomy	88·6	19·9	15·4	77·4	19·9	15·4	—	75·1	9·9	11·4	5	M	1	115·1
5	3·7	2·3	M	86·8	18·5	14		18·5	14·3	77·3	78·7	*6·3*	11	*4*	Pronounced	1	*174·6*
6	2·6	1·3	R	89·6	18·3	14·5		18·6	14·4	77·4	78·3	10·5	10·5	5	M	2	100
7	3	2·4	R	91·8	19	14·75		19·2	14·65	77·6	71·1	10·95	12·6	5	S	4	115·1
8	3	3·5	N	81·8	20·4	15·5[a]	76	20·4[a]	15·5[a]	—	73·9	10·1	12·25	5	S	3	121·2
9	3	3·4	N	81·8	18·8	14[a]	–	18·8	15	79·8	73·7	9·7	12·1	5	S	1	124·7
10	3·6	2·65	R	90·8	19·7	13·5	68·5	19·4	13·8	-	74·4	9·3	11·9	5	S	1	127·9
11	4·7	1·7	M	75	19·9	16	80·4	19·9	15·6	-	72·4	8·6	10·95	5	P	1	127·3
12	2·75	2·2[a]	M	82·5	19	14·4	75·8	19	13·75		74·1	10·25	11·1	5	S	4	108·3
13	3	?	R	87	18	14·5[a]		18·6	14·5[a]	72·6	70·9	9·1	12·2	5	M	2	134·1
14	3·75	3	M	79·6	19·1	14·8[a]		19·8	14·5	73·2	82·5	9·8	11·9	5	M	3	121·4
15	3·25	2·8	M-R	85·9	19·1	14·6	76·5	19·1	14·4	—	73·6	10·6	12	*6*	M	1	113·2
16	3	3·25	R	90	19	15·5[a]	76·3	19	14·5	—	76·8	9·65	11·75	5	M	1	121·7
17	3	2·75	M	80	20	15·2	—	20·5	15·2[a]	74·1	75·9	11·3	12·2	5	M	1	107·9
18	3	2·5	M	89·7	18·6	14·4	77·4	18·6[a]	14·4[a]	—	72·9	8·1	12·3	5	M	1	151·8
19	3·5	2·7	M	75	18·4	14·6		18·7	14·4	77	71·2	8·4	12·3	5	M	1	146·4
20	3·25	2·35	M	77·5	19·2	14·8	77·1	19·2	14·6	-	77·6	9·1	11·1	5	P	1	121·9
21	4	2·7	R	108·8	19·8	14·4	72·7	19·7	14·5	—	73·3	9·6	12·6	5	M	1	131·2
22	3·25	3[a]	M	*108·9*	19·5	14·2[a]	72·8	19·5[a]	14·2	–	72·2	10·6	12·5	5	S	2	117·9
23	3	3	M	91·1	19	15	—	19·2	15	78·1	70·5	8·9	12·1	5	M	1	135·9
24	3·5	3·6[a]	M	75·8	18·6	13·7	—	19·1	14·1	73·8	75·6	10·5	10·85	5	M	1	103·3
25	3·5	2	R	90·8	17·8	14	78·7	17·8	14		70·5	9·9	10·8	5	S	3	109·1
26	3	3·1	M	73·1	20·8	15·5	—	21·2	15·5	72·1	74·4	11·6	12·5	5	S	2	107·7
27	3·75	2·9	M	75·2	19·6	15·3	78·1	19·6	15·2	—	71·3	8·4	11·9	5	M	1	141·6
28	2·6	2·5	M	86·8	19·1	14	—	19·7	13·7	69·5	83·8	9·9	11·2	5	M	3	113·1
29	[illegible]	3·3	R	91·8	21·1	15·1	—	21·4	15·1[a]	70·5	79·3	11·25	12·7	5	M	2	112·9
30	3	2·8	R	84·1	19·4	15·2	78·8	19·4	15·2[a]	—	76·1	12·9	11·1	*6*	S	5	*86*
31	2·7	3·5	M-R	94·5	18·4	14·5	73·3	18·4	14·5	—	70·8	*13·2*[a]	11·4	*6*	S	3	86·3
32	3·5	2·9	N M	78·6	19·2	13·8[a]	71·3	19	13·8	—	73·8	10[a]	11·7	5	M	1	117
33	[illegible]	3[a]	N	73·6	18	12	—	18·2	13·8	75·8	79·1	10·6	10·3	5	S	1	97·1
34	4·35	3·4	N	66·6	20·5	16	78	20·5	15·8	—	71·6	10·2	12·2	5	M	3	119·6
35	3·7	2·5	M	90·9	18·7	13·8	—	18·8	14·1	75	75·2	9·5	11·4	5	M	4	120
36	4	2·6	R	90	18·6	15		18·8	14·5	77·2	75·2	11·1	10·9	5	S	1	98·2
37	3·75	3·2	M	79·2	*22·2*	16·5		*22·5*	16·5[a]	73·3	81·8	11·1	12·6	5	P	1	113·5
38	3·5[a]	3·1	M	84·9	19	13·5[a]	—	19·2	13·5	70·3	73·8	10·1	11·2	5/6	S	3	110·9
39	3·5	2·8	M	88	20·2	15·2	—	20·7	15·2[a]	73·4	76·7	10·2	12·2	5	M	2	119·6
40	3	2·6	R	100	21	15·5	—	21·3	15·6	73·2	81·1	11·1	12	5	M	3	108·1
41	3·5	2·9	N	85·8	18	13[a]	72·2	18[a]	13[a]	—	78·2	8·8	11·7	5	M	4	133
42	3·75	2·7	M	85	19·8	15·4	77·8	19·8	15·2	—	76·9	12·1	11·3	*6*	M	4	93·3
43	3·5	3·3	R	86·2	21·3	16	75·1	21·2	16	—	76·1	11	13·7	5	M	3	124·5
44	3	2·6	N	79·6	19·5	14·5	73·8	19·5	14·5	—	84·4	10·35	10·25	5	S	2	99
45	4	2·7	M	87·7	21·0	15·6	--	21·2	15·75	74·3	82·3	7·7	12·2	5	P	1	158·4
46	3·5	3·5	N	83·3	20·1	15	74·6	20	15·1	—	78·8	9·5	10·8	5	M	3	113·7
47	3·7	3·0	N	68·9	19·4	13·2	68	19·1	13·7	—	73·2	10·9	11·45	5	S	4	105

1. One sacral segment probably from coccygeal end. 3. Last lumbar articulates with left sacral ala. 5. Four sacral segments. No evidence of more at either end. 6. First sacral transitional. False promontory. 7. False promontory. 10. Congenital dislocation of right hip. Two sockets. Inlet diagonals, right 12·3, left 12·3. 11. Coccyx ossified to sacrum (not included), distorted. 12. False promontory. 13. Wide outlet (Plate VII, 9). 15. Well proportioned symmetrical sacrum (Plate VII, 8 and 10). 17. Coccyx ossified to sacrum (not included). 20. Tapering sacrum.

(Note: "a" means an approximate measurement often due to a broken bone. Italics signify high or low value or some peculiarity referred to in text.)

21. Ossification, left ilio-sacral synchondrosis. 26. Large pelvis, heavy bones. 30. Six sac. vert., one from coccyx (?). False promontory. 31. Six sac. vert., one from lumbar. False promontory. 33. Small pelvis, coccyx ossified to sacrum. 34. Sacral angle acute (3rd vertebra) (Plate VII, 9). 35. Narrow flat inlet. 37. Large pelvis. 38. Left sacral ala rises to 3·1 cms. above body and articulates with last lumbar. 42. Six sac. vert., one from lumbar(?). False promontory. 44. Small pelvis. "Male type." 46. "Male type" (Plate III, 3). 47. "Male type"

Table of Pelvic Measurements of

No.	U. S. N. M. No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Inlet: Oblique Diameter	Inlet: Obstetric Conjugate	Inlet: Difference	Inlet: Index	Pelvic: Antero-posterior Diameter	Pelvic: Fetal Head	Pelvic: Inter-tubers	Pelvic: Posterior Sagittal	Pelvic: Pubes to Perpendicular
48	213,331	Arizona	26·1	22·7	12·8	11	9·3	1·7	72·6	10·6	9·5	9·5	6·75	5·3
49	229,363	,,	28·2	26·0	*14·7*	12·2	10·2	2	69·4	13	13[a]	*12·7*[a]	8·5	5·7
50	239,201	,,	27·5	25	14·3	12·2	10·6	1·6	74·1	12·6	12	10·9	8·2	6
51	239,202	,,	26	22·5	13·3	11·1	8·9	2·2	67	12·3	11	10·5	6·6	6·5
52	239,203	,,	28	23·5	13	12·1	10·2	1·9	78·5	11·4	11	10·2	8·2	4·85
53	239,204	,,	25·75	23·5	12·4	10·4	9·3	1·1	75	12·6	10·5	9·2	8	6·2
54	239,215	,,	25	22·25	13	12·2	10·25	2	78·5	10·9	11·5	10·7	8	5·8
55	239,291	,,	25·25	22·25	12·25	12·3	10·85	1·5	88·5	10·5	10·5	9·2	7·5	4·4
56	239,293	,,	26·25	24·75	13·2	12·4	10·7	1·7	81	12	11·5	10·7	8·1	5
57	239,298	,,	25·5	20·5	13·1	12·5	10	2·5	76·3	10·9	12	10·9	8·3	3·75
58	239,305	,,	25·75	22	13·4	10·8	9·9	0·9	73·9	11·25	11·5	10·3	8·7	4·5
59	239,309	,,	27[a]	22·75[a]	12·3	13·2	11·5[a]	1·7	93·7	*12·5*	10[a]	*8*[a]	7·5	5·7
60	239,318	,,	23	21·25	12·3	10·1	8·5	1·6	69·1	10·25	11	9·7[a]	8·3	2·9
61	239,333	,,	28·5	24·75	14·2	12·1	10·1	2	71·1	11·7	12	11·2	7·6	5·4
62	239,348	,,	*20·8*	*17·2*	*10·7*	10·5	8·5	2	79·4	10·25	9·3	8·3	6·8	4·6
63	226,292	,,	26	22·5	13	11	9·2	1·8	70·7	11·4	11	9·8	7·5	4·8
64	239,385	,,	24·25	21·5	12·3	12·3	9·1	3·2	74	10·6	10·5	9·6	6·6	5
65	239,446	,,	25·75	20	13·6	11·4	9	2·4	66·2	10·5	10·5	10·4	6·2	5·8
66	239,453	,,	27·5	24·25	13·7	11·3	10	1·3	73	12·3	11·5	10·3	7·7	5·5
67	239,474	,,	26	24	13	9·9	*8*[a]	1·9	*61·5*	11	9·5	8·5	6·7	5·2
68	255,129	Arkansas	27·5	25·5	13·5	13·8	11·5	2·3	85·2	12·9	12·5	10·6	9·1	3·7
69	258,768	,,	26·5[a]	21·5	12·1	12·8	11	1·8	91	11·5	11	9·8	8	5·5
70	259,301	,,	25·15	21·8	13·6	12·8	11	1·8	80·9	11	12	11·4	7·6	3·9
71	225,253	Choctaw	26·9	24·5	12·6	13	10·7	2·3	84·6	11·9	10	9	6·5	6·6
72	49,735	Colorado	26·75	23·75	13·3	11·2	8·7	2·5	65·4	11	12	11·4	7	4·4
73	225,214	Comanche	26·1	23·4	13	11·1	8·3	2·8	63·8	9·4	*9*	8·5	6·6	3·7
74	248,579	Eskimo (St Michaels)	27·75	21	13·6	14·3	11·8	2·5	86·8	13·4	10·25	8·75	8·25	6·5
	Peabody M.													
75	12,804	Tennessee	25	22·3	12·8	13·5	11·5	2	89·8	11·4[a]	11	10·2	7·6[a]	4·6
	U. S. N. M.													
76	227,434	Illinois	22	21	12·2	10	8·6	1·4	70·5	10·8	9·5	9[a]	7	5·3
77	227,440	,,	24·5	22[a]	12·5	11·6	9·4	2·2	75·2	11·8	10	9·35	7	5·9
78	227,441	,,	21·9	17·25	11·5	11	9·7	1·3	84·3	11	10·5	8·85	8·1	4
79	227,445	,,	24·35	20·5	13·3	14	12[a]	2	90·2	12	11	10·3	8·8	5·2
80	227,448	,,	25	21·75[a]	12·1	13·6	11·8	1·8	97·5	11·5	10	9·1	7	6·1
81	227,450	,,	23·75	20·75[a]	12·8	13·3	11·75	1·6	91·8	11·5	10·5	8·85	8	5·1
82	225,420	Kentucky	26·8	23·8	12·8	11·6	10·1	1·5	78·9	11·8	9·5	8·5	7	6·5
83	225,422	,,	28·3	25·7	13·4	11·1	9·8	1·3	73·1	12·1	11[a]	9	8[a]	6·9
84	225,421	,,	24·2	20·5	12·6	11·7	10·4	1·3	82·8	11·9	11·5	10·85	8·7	4·5
85	225,423	,,	26·3	23·9	12·6	11·9	10·2	1·7	81·3	10·6	9·5	8·6	6·9	6·4
86	225,425	,,	26·5	22·5[a]	12·8	13·5	12	1·5	93·3	13·1	12	10·8	8·1	5·8
87	255,105	Louisiana	28·5[a]	25[a]	13·3	11	9·2	1·8	69·2	11	10·5	9·8[a]	6·7[a]	6·3
88	255,214	,,	26·5	24	13·1	13·1	11·3	1·8	86·2	10·9	10·5	10	7·2	5·2
89	255,216	,,	25·75	22	13·8	12·4	10·6	1·8	76·8	12·2	11·25	10	8·1	6·1
90	216,213	Mexico	24·5[a]	19·5	11·5	13·2	11·1	2·1	96·1	11·2	10·75	10·2	7·1	5·1
91	228,925	New Mexico	26[a]	24	13·5	12·2	10·8	1·4	80	11	11	10·35	7·5	4·6
92	228,950	,,	26[a]	21·5	12·4	11·9	10	1·9	80·6	12·3	11	10	7·3	6
93	228,967	,,	27[a]	24[a]	13	12·8	10·4	2·4	80	12	10·25	9·1	7·5	5·3

51. (Plate IV, 4) sacrum. Transitional vertebra. 53. Six sac. vert., one from coccyx (?). 57. Double promontory (Plate IV, 4). 59. Narrow arch. Exostoses on sacrum. 60. Four sac. vert. Extreme "female type" (Plate IV, 4). 62. Pygmy or very small pelvis (Plate I, 1). 67. Small pelvis. 68. Six sac. vert., one from coccyx (?). 71. Coccyx ankylosed to sacrum (not included). 72. Spinal canal open from 1st sacral down. 73. Rather flat pelvis (Plate II, 2). 74. Six sac. vert., one from

American Indian Squaws (continued).

Outlet				Innominate Bone						Pelvic Index	Sacrum					
Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index		Height	Breadth	Number Segments	Curve	Segment begins at	Index
3	2·5	N-M	89·6	18·4	14·2	—	18·6	13·9	72·2	71·2	9·7	10·9	5	M	1	112·4
3·5	3·4	R	90·8	20·4[a]	15·5	—	20·4	15·5	75·9	72·3	10·8	11·5[a]	5	S	4	106·5
3·75	2·7	R	86·5	19·3	15·5	—	19·4	15·5	79·9	70·5	9·6	11·6	5	M	2	120·8
3·5	3·5	M	85·3	18·7	14·1	—	19·1	14·1	73·8	73·4	11·4	10·8	5/4	S	2	94·7
4·25	3·2	R	89·5	20·2	15·4	76·4	20·2	15·2	—	72·1	9	11·5	5	M	1	127·8
3·25	2·4	R-M	73	18·7	13·9	79·7	18·7	13·8	—	72·6	10·9	10·9	*6*	S	3	100
3·25	2·7	R	98·1	18·5	13·4	—	18·5	13·7	74	74	11·7	10·9	5	M	1	92·3
3·25	2·6	R	87·6	18·8	13·9	—	19·1	13·6	71·2	75·6	9·9	10·6	5	M	3	107
3·5	3·2	M	89·2	18·5	13·8	74·6	18·5	13·8	—	70·4	9·9	11·3	5	S	2	114·1
3·5	2·6	R	100	19·6	14·2	—	19·8	13·7	69·2	77·6	10·9	11·3	5	M	2	103·7
4	2	R	92·4	19[a]	14	—	19	14	73·7	73·8	8	11·1	5	M	1	138·8
3	2·7	N	*64*	19·4	14·3	73·7	19·4[a]	14·3[a]	—	71·3	10·9	11·5	5	M	1	118·1
3·5	2·6	R	94·6	17·2	13	—	17·3	12·9	74·5	75·2	7·25	10·9[a]	*4*	M	1	150·3
3·5	3	R	94·9	19·8	15·3	—	19·9	15·4	77·4	69·8	10·5	12·4	5	M	2	118·1
2·8	2·7	M	80·9	*16*	*11·3*	70·6	*15·7*	*11·15*	—	76·9	7·9	*9·2*	5	S	3	116·5
3·5	2·5	M	85·1	18·3	13·5[a]	73·8	18·3	13·5[a]	—	70·4	9·4	11	5	S	3	117
2·75	3·2	M	95·6	18·8	14·8	78·7	18·7	14·7	—	77·5	9·7	10·7	5	M	2	110·3
4	3·1	M	99	19·4	13·9	—	19·5	13·9	70·8	75·7	9·4	12·1	5	M	1	128·7
4	2·8	N-M	87·7	18·7	14	—	18·7	14·1	75·4	68	8·5	11·6	5	M	2	136·5
3·25	3·6[a]	N	77·3	18·1	13·6[a]	—	18·1	13·6	75·1	69·6	8·2	10·8	5	M	1	131·7
3	3·3	R	82·2	20·8	15·25	—	21	15·5	73·8	76·4	10·8	12·7	*6*	M	2	117·6
3·75	3·5	M-R	85·2	19·2	14·6	—	19·3	14·6[a]	75·6	72·8	9·2	11·45	5	P	1	124·5
3	3·2	M	103·6	20·7	14·4[a]	—	20·7	14·4	69·5	82·3	10·3	11·7	5	M	1	113·6
3·25	3·5	M	75·6	18·2	15·3	*84·1*	18·1	15·2	—	67·7	10·9	11·5	5	S	1	105·5
2·5	?	R	103·6	19	14·1	—	19·1	14·1	73·8	73·6	10	11·2	5	M	1	112
3·25	3·2	M	90·4	19·1	14·4	75·8	19	14·1	—	72·8	11·8	11·3	5	M	2	95·8
3·5	3·8	N	65·3	21·9	15·7	71·3	21·7	15·1	—	78·9	10·9	11·4	*6*	M	2	104·6
3·7	2·2	R	89·5	18·8	14·1	—	19	13·8	72·6	76	10·9[a]	11·5	*6*	M	2	105·5
3·75	2·85	?	83·3	17·5	12·9	—	17·7	12·75	72	80·4	7·75	10·3	5	P	1	132·9
3·5	3·6	N-M	79·2	19·2	14	—	19·5	14[a]	79·6	75·5	8·2	11·3	*6*	P	3	137·8
3	2·9	N-M	80·4	18·4[a]	13·9	—	18·4	13·9[a]	70·1	84	7·8	10·5[a]	5	M	2	134·6
3·5	?	M	85·8	20·5	15·2	74·1	20·4	15·2[a]	—	84·2	11	11·5	5+?	M	3	104·5
3	3·4	N	79·1	18·6	14·5	—	18·6	14·5[a]	77·9	74·4	8·5	10·35	5	P	3	121·8
3·5	3·6	N	76·9	20·8	15·4	74	20·6	15·4	—	87·1	10·2	11·4	5	M	1	111·7
3·5	3	N	72·1	19·3	13·5	—	19·4	13·5	69·6	72·4	10·3	12·7	5	M	1	123·3
2·7	3·1	R	74·4	20[a]	15·5	77·5	20	15·3	—	70·7	12·5	11·7	5/6	S	4	93·7
2·5	2·6[a]	R	90·8	18	14·6[a]	—	18·6	14·6[a]	78·5	76·8	10·2[a]	11·7	5	M	3	114·7
3·75	3	M	81	19·2	14·7	76·5	19·2[a]	14·7	—	73	10·1[a]	11·2	5	S	1	110·9
3·5	3·5[a]	R	74·8	20	15·9	79·5	20[a]	15·9[a]	—	75·5	10·3	10·7	5	M	1	103·9
3·75[a]	3·2	M	89·1	18·1	14·7[a]	81·2	18·1[a]	14·7[a]	—	*63·7*	9[a]	12·1[a]	5	M	2	134·4
3	3	M	91·7	19·2	14·5	—	19·3	14·5[a]	75·1	72·8	10·5	12	*6*	M	2	114·3
4	3·3	M	81·9	20	15·15	75·7	19·9	15·4	—	77·6	11·5	12	*6*	M	2	104·4
3·25	3·4	R	91·1	18·9	14	—	19·2	13·8	71·8	78·4	10·8	10·5	5	M	3	97·2
3·25	3·1	M	94·1	19·4	13·8[a]	71·1	19·4[a]	13·8	—	74·6	10·9	11·6	5	M	3	106·4
3·4	3·1	M	91·3	18·9	13·65	72·2	18·9	13·6[a]	—	72·7	8·4	11·1	5	M	2	132·1
3	3	M	75·8	19·3	13·1[a]	*67·8*	19·3[a]	13·1[a]	—	71·5	9·7	11·1	5	M	2	114·4

lumbar(?) (Plate IV, 4). 75. Six sac. vert., all sacral in character. 77. (Plate V, 5), six sac. vert., one from coccyx. 1st lumbar forms false promontory. 78. Small pelvis. 79. Suggestive of six sac. vert., one from coccyx. 82. Coccyx ossified to sacrum. 83. 5th lumbar sacralized left, free, right. False promontory. 86. Round pelvic inlet (Plate II, 2). 88. Six sac. vert., one from coccyx (?). 89. Six sac. vert., one from coccyx(?). 92. Coccyx ossified to sacrum.

Table of Pelvic Measurements of American Indian Squaws (continued).

No.	U.S.N.M. No.	Tribe or Locality	Inter-crests cms	Inter-spines	INLET Transverse	Oblique Diameter	Obstetric Conjugate	Difference	Index	PELVIC Antero-posterior Diameter	Fetal Head	Inter-tubers	Posterior Sagittal	Pubes to Perpendicular
48	213,331	Arizona	26·1	22·7	12·8	11	9·3	1·7	72·6	10·6	9·5	9·5	6·75	5·3
49	229,363	"	28·2	26·0	*14·7*	12·2	10·2	2	69·4	13	13[a]	*12·7*[a]	8·5	5·7
50	239,201	"	27·5	25	14·3	12·2	10·6	1·6	74·1	12·6	12	10·9	8·2	6
51	239,202	"	26	22·5	13·3	11·1	8·9	2·2	67	12·3	11	10·5	6·6	6·5
52	239,203	"	28	23·5	13	12·1	10·2	1·9	78·5	11·4	11	10·2	8·2	4·85
53	239,204	"	25·75	23·5	12·1	10·4	9·3	1·1	75	12·6	10·5	9·2	8	6·2
54	239,215	"	25	22·25	13	12·2	10·25	2	78·5	10·9	11·5	10·7	8	5·8
55	239,291	"	25·25	22·25	12·25	12·3	10·85	1·5	88·5	10·5	10·5	9·2	7·5	4·4
56	239,293	"	26·25	21·75	13·2	12·4	10·7	1·7	81	12	11·5	10·7	8·1	5
57	239,298	"	25·5	20·5	13·1	12·5	10	2·5	76·3	10·9	12	10·9	8·3	3·75
58	239,305	"	25·75	22	13·4	10·8	9·9	0·9	73·9	11·25	11·5	10·3	8·7	4·5
59	239,309	"	27[a]	22·75[a]	12·3	13·2	11·5[a]	1·7	93·7	*12·5*	10[a]	8[a]	7·5	5·7
60	239,318	"	23	21·25	12·3	10·1	8·5	1·6	69·1	10·25	11	9·7[a]	8·3	2·9
61	239,333	"	25·5	21·75	11·2	12·1	10·1	2	71·1	11·7	12	11·2	7·6	5·4
62	239,348	"	*20·8*	*17·2*	*10·7*	10·5	8·5	2	79·4	10·25	9·3	8·3	6·8	4·6
63	226,292	"	26	22·5	13	11	9·2	1·8	70·7	11·4	11	9·8	7·5	4·8
64	239,385	"	24·25	21·5	12·3	12·3	9·1	3·2	74	10·6	10·5	9·6	6·6	5
65	239,446	"	25·75	20	13·6	11·4	9	2·4	66·2	10·5	10·5	10·4	6·2	5·8
66	239,453	"	27·5	21·25	13·7	11·3	10	1·3	73	12·3	11·5	10·3	7·7	5·5
67	239,474	"	26	24	13	9·9	8[a]	1·9	*61·5*	11	9·5	8·5	6·7	5·2
68	255,129	Arkansas	27·5	25·5	13·5	13·8	11·5	2·3	85·2	12·9	12·5	10·6	9·1	3·7
69	258,768	"	26·5[a]	21·5	12·1	12·8	11	1·8	91	11·5	11	9·8	8	5·5
70	259,301	"	25·15	21·8	13·6	12·8	11	1·8	80·9	11	12	11·4	7·6	3·9
71	225,253	Choctaw	26·9	24·5	12·6	13	10·7	2·3	84·6	11·9	10	9	6·5	6·6
72	49,735	Colorado	26·75	23·75	13·3	11·2	8·7	2·5	65·4	11	12	11·4	7	4·4
73	227,241	Comanche	26·1	23·4	13	11·1	8·3	2·8	63·8	9·4	*9*	8·5	6·6	3·7
74	248,579	Eskimo (St Michaels)	27·75	21	13·6	14·3	11·8	2·5	86·8	13·4	10·25	8·75	8·25	6·5
	Peabody M.													
75	12,804	Tennessee	25	22·3	12·8	13·5	11·5	2	89·8	11·4[a]	11	10·2	7·6[a]	4·6
	U.S.N.M.													
76	227,434	Illinois	22	21	12·2	10	8·6	1·4	70·5	10·8	9·5	9[a]	7	5·3
77	227,440	"	24·5	22[a]	12·5	11·6	9·4	2·2	75·2	11·8	10	9·35	7	5·9
78	227,441	"	21·9	17·25	11·5	11	9·7	1·3	84·3	11	10·5	8·85	8·1	4
79	227,445	"	21·35	20·5	13·3	14	12[a]	2	90·2	12	11	10·3	8·8	5·2
80	227,448	"	25	21·75[a]	12·1	13·6	11·8	1·8	97·5	11·5	10	9·1	7	6·1
81	227,450	"	23·75	20·75[a]	12·8	13·3	11·75	1·6	91·8	11·5	10·5	8·85	8	5·1
82	225,420	Kentucky	26·8	23·8	12·8	11·6	10·1	1·5	78·9	11·8	9·5	8·5	7	6·5
83	225,422	"	28·3	25·7	13·4	11·1	9·8	1·3	73·1	12·1	11[a]	9	8[a]	6·9
84	225,421	"	24·2	20·5	12·6	11·7	10·4	1·3	82·8	11·9	11·5	10·85	8·7	4·5
85	225,423	"	26·3	23·9	12·6	11·9	10·2	1·7	81·3	10·6	9·5	8·6	6·9	6·4
86	225,425	"	26·5	22·5[a]	12·8	13·5	12	1·5	93·3	13·1	12	10·8	8·1	5·8
87	255,105	Louisiana	28·5[a]	25[a]	13·3	11	9·2	1·8	69·2	11	10·5	9·8[a]	6·7[a]	6·3
88	255,214	"	26·5	24	13·1	13·1	11·3	1·8	86·2	10·9	10·5	10	7·2	5·2
89	255,216	"	25·75	22	13·8	12·4	10·6	1·8	76·8	12·2	11·25	10	8·1	6·1
90	246,213	Mexico	24·5[a]	19·5	11·5	13·2	11·1	2·1	96·1	11·2	10·75	10·2	7·1	5·1
91	228,925	New Mexico	26[a]	24	13·5	12·2	10·8	1·4	80	11	11	10·35	7·5	4·6
92	228,950	"	26[a]	21·5	12·4	11·9	10	1·9	80·6	12·3	11	10	7·3	6
93	228,967	"	27[a]	24[a]	13	12·8	10·4	2·4	80	12	10·25	9·1	7·5	5·3

No.	OUTLET Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	INNOMINATE BONE Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index	Pelvic Index	SACRUM Height	Breadth	Number Segments	Curve	Segment begins at	Index
48	3	2·5	N-M	89·6	18·4	14·2		18·6	13·9	72·2	71·2	9·7	10·9	5	M	1	112·4
49	3·5	3·4	R	90·8	20·4[a]	15·5	—	20·4	15·5	75·9	72·3	10·8	11·5[a]	5	S	4	106·5
50	3·75	2·7	R	86·5	19·3	15·5		19·4	15·5	79·9	70·5	9·6	11·6	5	M	2	120·8
51	3·5	3·5	M	85·3	18·7	14·1	—	19·1	14·1	73·8	73·4	11·4	10·8	5/4	S	2	94·7
52	4·25	3·2	R	89·5	20·2	15·4	76·4	20·2	15·2		72·1	9	11·5	5	M	1	127·8
53	3·25	2·4	R-M	73	18·7	13·9	79·7	18·7	13·8		72·6	10·9	10·9	6	S	3	100
54	3·25	2·7	R	98·1	18·5	13·4	—	18·5	13·7	74	74	11·7	10·9	5	M	1	92·3
55	3·25	2·6	R	87·6	18·8	13·9		19·1	13·6	71·2	75·6	9·9	10·6	5	M	3	107
56	3·5	3·2	M	89·2	18·5	13·8	74·6	18·5	13·8	.	70·4	9·3	11·3	5	S	2	111·1
57	3·5	2·6	R	100	19·6	14·2		19·8	13·7	69·2	77·6	10·9	11·3	5	M	2	103·7
58	4	2	R	92·4	19[a]	14	–	19	14	73·7	73·8	8	11·1	5	M	1	138·8
59	3	2·7	N	*64*	19·4	14·3	73·7	19·4[a]	14·3[a]	—	71·3	10·9	11·5	5	M	1	148·1
60	3·5	2·6	R	94·6	17·2	13		17·3	12·9	74·5	75·2	7·25	10·9[a]	*4*	M	1	150·3
61	3·5	3	R	94·9	19·8	15·3	—	19·9	15·4	77·4	69·8	10·5	12·4	5	M	2	118·1
62	2·8	2·7	M	80·9	*16*	*11·3*	70·6	*15·7*	*11·15*	—	76·9	7·9	*9·2*	5	S	3	116·5
63	3·5	2·5	M	85·1	18·3	13·5[a]	73·8	18·3	13·5[a]		70·4	9·4	11	5	S	3	117
64	2·75	3·2	M	95·6	18·8	14·8	78·7	18·7	14·7		77·5	9·7	10·7	5	M	2	110·3
65	4	3·1	M	99	19·4	13·9		19·5	13·9	70·8	75·7	9·4	12·1	5	M	1	128·7
66	4	2·8	N-M	87·7	18·7	14		18·7	14·1	75·4	68	8·5	11·6	5	M	2	136·5
67	3·25	3·6[a]	N	77·3	18·1	13·6[a]		18·1	13·6	75·1	69·6	8·2	10·8	5	M	1	131·7
68	3	3·3	R	82·2	20·8	15·25		21	15·5	73·8	76·4	10·8	12·7	6	M	2	117·6
69	3·75	3·5	M-R	85·2	19·2	14·6		19·3	14·6[a]	75·6	72·8	9·2	11·45	5	P	1	124·5
70	3	3·2	M	103·6	20·7	14·4[a]		20·7	14·4	69·5	82·3	10·3	11·7	5	M	1	113·6
71	3·25	3·5	M	75·6	18·2	15·3	*84·1*	18·1	15·2		67·7	10·9	11·5	5	S	1	105·5
72	?	?	R	103·6	19	14·1		19·1	14·1	73·8	73·6	10	11·2	5	M	1	112
73	3·25	3·2	M	90·4	19·1	14·4	75·8	19	14·1	—	72·8	11·8	11·3	5	M	2	95·8
74	3·5	3·8	N	65·3	21·9	15·7	71·3	21·7	15·1	—	78·9	10·9	11·4	*6*	M	2	104·6
75	3·7	2·2	R	89·5	18·8	14·1	—	19	13·8	72·6	76	10·9[a]	11·5	*6*	M	2	105·5
76	3·75	2·85	?	83·3	17·5	12·9		17·7	12·75	72	80·4	7·75	10·3	5	P	1	132·9
77	3·5	3·6	N-M	79·2	19·2	14		19·5	14[a]	79·6	75·5	8·2	11·3	*6*	P	3	137·8
78	3	2·9	N-M	80·4	18·4[a]	13·9		18·4	13·9[a]	70·1	84	7·8	10·5[a]	5	M	2	134·6
79	3·5	?	M	85·8	20·5	15·2	71·1	20·4	15·2[a]		84·2	11	11·5	5+?	M	3	104·5
80	3	3·4	N	79·1	18·6	14·5		18·6	14·5[a]	77·9	74·4	8·5	10·35	5	P	3	121·8
81	3·5	3·6	N	76·9	20·8	15·4	74	20·6	15·4	-	87·1	10·2	11·4	5	M	1	111·7
82	3·5	3	N	72·1	19·3	13·5	.	19·4	13·5	69·6	72·4	10·3	12·7	5	M	1	123·3
83	2·7	3·1	R	74·4	20[a]	15·5	77·5	20	15·3	—	70·7	12·5	11·7	5 6	S	4	93·7
84	2·5	2·6[a]	R	90·8	18	14·6[a]	—	18·6	14·6[a]	78·5	76·8	10·2[a]	11·7	5	M	3	114·7
85	3·75	3	M	81	19·2	14·7	76·5	19·2[a]	14·7	-	73	10·1[a]	11·2	5	S	1	110·9
86	3·5	3·5[a]	R	74·8	20	15·9	79·5	20[a]	15·9[a]	—	75·5	10·3	10·7	5	M	1	103·9
87	3·75[a]	3·2	M	89·1	18·1	14·7[a]	81·2	18·1[a]	14·7[a]	—	65·7	9[a]	12·1[a]	5	M	2	134·4
88	3	3	M	91·7	19·2	14·5	—	19·3	14·5[a]	75·1	72·8	10·5	12	*6*	M	2	111·3
89	4	3·3	M	81·9	20	15·15	75·7	19·9	15·4		77·6	11·5	12	*6*	M	2	104·4
90	3·25	3·4	R	91·1	18·9	14	..	19·2	13·8	71·8	78·4	10·8	10·5	5	M	3	97·2
91	3·25	3·1	M	94·1	19·4	13·8[a]	71·1	19·4[a]	13·8	-	74·6	10·9	11·6	5	M	3	106·4
92	3·4	3·1	M	91·3	18·9	13·65	72·2	18·9	13·6[a]	-	72·7	8·4	11·1	5	M	2	132·1
93	3	3	M	75·8	19·3	13·1[a]	*67·8*	19·3[a]	13·1[a]	—	71·5	9·7	11·1	5	M	2	114·4

51. (Plate IV, 4) sacrum. Transitional vertebra. 53. Six sac. vert., one from coccyx (?). 57. Double promontory (Plate IV, 4). 59. Narrow arch. Exostoses on sacrum. 60. Four sac. vert. Extreme "female type" (Plate IV, 4). 62. Pygmy or very small pelvis (Plate I, 1). 67. Small pelvis. 68. Six sac. vert., one from coccyx (?). 71. Coccyx ankylosed to sacrum (not included). 72. Spinal canal open from 1st sacral down. 73. Rather flat pelvis (Plate II, 2). 74. Six sac. vert., one from lumbar (?) (Plate IV, 4). 75. Six sac. vert., all sacral in character. 77. (Plate V, 5), six sac. vert., one from coccyx. 1st lumbar forms false promontory. 78. Small pelvis. 79. Suggestive of six sac. vert., one from coccyx. 82. Coccyx ossified to sacrum. 83. 5th lumbar sacralized left, free, right. False promontory. 86. Round pelvic inlet (Plate II, 2). 88. Six sac. vert., one from coccyx (?). 89. Six sac. vert., one from coccyx (?). 92. Coccyx ossified to sacrum.

Table of Pelvic Measurements of

No.	U. S. N. M. No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Inlet: Oblique Diameter	Inlet: Obstetric Conjugate	Inlet: Difference	Inlet: Index	Pelvic: Antero-posterior Diameter	Pelvic: Fetal Head	Pelvic: Inter-tubers	Pelvic: Posterior Sagittal	Pelvic: Pubes to Perpendicular
94	246,968	New Mexico	27·5	23·8	14·4	13·3	10·8	2·5	75	12·6	11	10	8	5·75
95	225,215	Pah Ute	25·5	22·2	12·7	12·5	10·5	2	82·7	13·5	14	11·1	11·7	2·8
96	204,254	Peru	27·3	21·5	13·8	12·2	9·25	2·9	67	11·3	10·25	9	7·5	5·3
97	225,213	„	26·5	22·25	12·5	10·4	8[a]	2·4	64	11·9	10	8·6	8·1	5·5
98	225,250	„	25	20·5	12·9	12	9·5	2·5	73·6	12·5	11	10·3	7·7[a]	6
99	227,006	„	24·6	20	12	12·3	10·8	1·5	90	11·1	10	9·6	6·3	5·7
100	227,007	„	27	24·5	13·5	14	11·6	2·4	86	11·75	12	11·75	7·4	5·7
101	227,011	„	25·3	23·3	12·3	12·1	10·5	1·6	85·4	12·5[a]	11[a]	9·5[a]	8·4[a]	4·1[a]
102	227,014	„	26	22	13·2	12·8	11·25	1·6	85·2	11·3[a]	11[a]	10	7[a]	5·5
103	227,024	„	27·3	26·5	14·3	13·1	11·35	1·8	79·4	13·6	12·5	11·5	8·8	6·1
104	227,432	„	24·5	22·3	12	11·4	9[a]	2·4	75	11·5	9·5	8·1	7	5·6
105	242,513	„	25	21·5	11·5	11·2	9·1	2·1	79·1	11·5[a]	11·5[a]	11	7·2[a]	5·3
106	242,568	„	24·5	22	11·8	11·6	9·4	2·2	79·7	11·1	9	8·3	6·3	6·4
107	225,217	Sioux	27·1	23	13·1	13·2	11·6	1·6	88·5	11·4	12	10·7	8·5	3·9
108	225,218	„	28·5	25	13·9	13·3	10·1	3·2	72·7	11·8	11	9·7	8	5·1
109	225,221	„	27·8	24·1	12·1	13·3	11·1	2·2	91·7	11	11	9·8	7·6	4·7
110	225,242	„	28	25·5	12·6	13·2	11·2	2	88·8	12·4	11·5	9·7	8·4	5·1
111	225,261	„	27	23·5	13	12·3	10	2·3	76·9	10·8	11	9·5	7·2[a]	4·1
112	225,415	„	29·1	25	14·1	14·4	12	2·4	85·1	13·2	11·25	10·6	7·6	6·5
113	225,416	„	28	24·1	12·75	11·8	10·15	1·7	79·6	12·5	10·5	9·4	7·75	6·4
114	169,672	Virginia	28[a]	24·5[a]	13·2[a]	11·8	10·2[a]	1·6	77·3	11·1	11	9·9	7·2	4·5
115	98,470	Wisconsin	25·5	22·6	13·25	?	10·85	?	81·9	10·8	10·3	9·25	7·15	5·3
	A. M. N. H., N.Y.													
116	H, 3658	Pueblo, N.M.	27	23·25	13	13·3	11·2	2·1	86·2	12·7	12·5	11	9·2	5
117	„	„	24·5	21·8	12·5	12·2	11	1·2	88	10·6	10·5	10·2	6·5	5·5
118	„	„	25·2	23	13·4	12·5	10·7	1·8	79·9	11·8	12·5	11·1	7·6	5·3
119	„	„	24·9	23[a]	13·1	11·4	9·5	1·9	72·5	12	10	9	7·4	5·9
120	„	„	27·3	25·4	13·3	12·3	10·4	1·9	77·4	12	10·5	8·5	7·5	5·2
121	H, 102	South Utah	23·8	19·3	12·4	11	9	2	72·6	11·6	12·5	11	9·4	4
122	H, 191	„	22	20	12	11·8	10·4	1·4	86·7	11·3	11	10·4	7·7	4·8
123	H, 271	„	26·5	23	13·1	11·8	10·1	1·7	77·1	11·8	10	9·6	7·2	5·8
124	H, 307	„	27	23·4	14	12·1	10·6	1·5	75·7	11·6	10·5	9·7	7·1	6·2
125	H, 321	„	25·7	22·9	13·1	10·2	9	1·2	68·7	10·3	10·5	9·7	6·8	4·7
126	99/2393	Mexico	24·5	22	12	11·6	9·9	1·7	82·5	11·2	10·5	9·5	7·1	5·7
127	99/2230	„	25	20·5	12·9	13·3	11·8	1·5	91·5	11·4	10	9	7	5·5
128	99/2165	„	26	23·7	12·5	11·5	10	1·5	80	11	10·5	8·9	7·5	5
129	99/305	„	25·6	22	13·2	11·2	9·6	1·6	72·7	10·3	10	9·2	6·6	4·6
130	99/2506	„	27	24·1	12·1	13·1	11·5	1·6	95	11·8	9·5	8·2	7·3	6
131	99/2426	„	26·9	24·6	13·1	12·1	10·5	1·6	80·1	10·2	9·5	9·2	6·1	5·3
132	99/2542	„	25·6	22	12·8	13·2	11·9	1·3	93	11·2	10	9·2	6·7	5·4
133	99/2330	„	25·1	23·5	13	14·9	13	1·9	100	12·6	11	9·5	8·5	5·3
134	99/2325	„	25·1	21·4	13	12·8	11·3	1·5	86·9	11·6	11	10	7·7	5·2
135	99/2539	„	28	24	13·5	12·8	11	1·8	81·5	12·7	11	9·1	8·1	5·8
136	99/2329	„	25·1	23·1	13·1	13·7	11·6	2·1	88·5	12·5	11	9·4	8·3	4·5
137	99/2544	„	26·1	21	12·5	12·9	11	1·9	88	10·3	10·5	9·7	8	4·4
138	99/2486	„	23·2	22	12·6	11·9	10	1·9	79·4	11·5	10	9	7	4·6

94. Six sac. vert. (Plate IV, 4). False promontory. Spinal canal of sacrum open. 95. Shallow pelvis, wide outlet (Plate III, 3). 98. Six sac. vert., one from coccyx (?). 99. Six sac. vert., all sacral in character (Plate IV, 4). 100. False promontory (Plate VI, 6). 102. Six sac. vert., one from coccyx (?). 104. Six sac. vert., one from coccyx (?). 105. Last sacral segment missing. 106. Transitional vertebra. No false promontory. Small pelvic cavity (Plate IV, 4, Plate VI, 7). 107. Large pelvis (Plate I, 1). 108. Large pelvis. 109. Transitional vertebra, left = 6, right = 5 segments (Plate IV, 4). 111. Coccyx ossified to sacrum (not included). 116. Wide outlet (bones dark

American Indian Squaws (continued).

Outlet				Innominate Bone						Pelvic Index	Sacrum					
Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index		Height	Breadth	Number Segments	Curve	Segment begins at	Index
3·25	3	M	79·4	20·8	15·35	73·8	20·7	15·25	—	72	11·9	11·9	*6*	P	2	100
3·75	3·1	R	82·2	19·8	14·8	—	19·9	14·9	74·9	77·6	7·5	11·6	5	P	1	154·7
4·25	3·15	N	79·6	20·7	15·2	—	20·7	15·3	73·9	76·2	10·7	12·4	5	M	3	115·9
3·25	2·4	N	72·3	18·4	14·5	73·4	18·3	14·2	—	69·4	8·85	10·7	5	M	2	120·9
3	2·2	M	82·4	19·8	15	—	20	15·1	75·5	80	10·5[a]	11·9	*6*	S	2	113·3
3	3·3	M	86·5	18·7	14·4	77	18·4	14·4	—	76	9·7	10·5	*6*	M	1	108·2
3·25	3·2	M	100	20·1	14·5	—	20·2	14·8[a]	73·3	74·8	10·6	11·7	5	M	3	110·4
3·25[a]	3[a]	M	76	18·2	14	76·9	18·2	14	—	70·2	9[a]	11·5	5	M	1	127·8
3·25	2·8	M	88·5	18·7	14·6	—	18·7	14·7	78·6	71·9	9·4[a]	11·5	*6*	S	1	111·7
3·25	3·5	R	87·4	20·2	15·6	—	20·4	15·1	74	74·7	10	11·7	5	M	1	117
3·1	4	N	70·4	18·6	13·75	73·9	18·3	13·75	—	75·9	9	11·2	*6*	M	2	124·4
3·25	3	R	95·7	17·4	13·25	—	17·4[a]	13 6	78·1	69·6	8·7[a]	10	5	M	1	114·9
2·75	3	M	74·8	18	14·1	78·3	18	14·1	—	73·5	9·8	10·5	5/6	M	2	107·7
3·75	3·1	R	93·9	20·4	16·1	—	20·5	16·1	78·5	75·6	9·9	13·2	5	M	1	133·3
4·75	3·4	N	82·3	21·3	16·6	—	21·4	15·8	73·8	75·1	9·6	11·3	5	P	1	117·7
3·5	3·3	N	89·1	20·1	15·6	77·6	21·1	15·5	—	75·9	10·8	11·5	5/6	M	3	106·6
4	3·2	N	78·2	20·2	15·5	—	20·5	15·3	74·6	73·2	10·5	12·1	5	S	1	115·2
3·25	2·9	M	87·9	20	15·1	75·5	19·9	15·1	—	74·1	9·8	11·5	5	M	1	117·3
3	3·7	N	80·3	21·5	*16·9*	78·6	21·5	*16·8*	—	73·8	12	12·15	5	M	2	101·2
3·4	3·8	N	75·2	20	16	80	19·4	15·85	—	71·4	8·1	12	5	P	1	148·1
2·75	3	M	89·2	18·8	15	—	19·1	15[a]	78·5	68·2	9·9	11·9	5	M	3	120·2
2·8	3·25	N	85·6	20	14·7	73·5	19·7	14·4	—	78·4	10·5	11·8	5	M	2	112·4
4·3	3·3	M	86·6	20·1	15·4	—	20·1	15·5	77·1	74·4	10·1	11·4	5	S	3	112·9
3·7	2·5	M	96·2	18·5	14·1	—	18·5	14·3	77·3	75·5	9·7	11·5	5	M	2	117·5
3·3	2	R	94·1	18·8	15·4	—	18·8	15·6	77·7	74·6	10·4	11·5	5	M	3	110·6
3·8	1·9	M	75	18·9	14·2[a]	—	19	14·2[a]	74·7	76·3	10[a]	11·9	5	M	1	119
3	1·6	M	70·8	18·2	14·7	—	19·1	15·2	79·6	70	11·3	11·7	5	S	2	103·5
4	2·6	R	94·8	18·3	13·3	72·7	18·2	13·3[a]	—	76·9	9·2	11	5	M	3	119·6
4	1·2	M	92	19·6	14·2	—	19·8	14·2	71·7	*90*	8·2	10·4	*4*	M	1	126·8
3·5	2·1	R	84·8	19·2	14·8	77·1	19·2	14·8	—	72·5	11·3	11·3	5	Reverse	5 (?)	100
4·1	2·2	M	83·6	20	14·3	—	20·3	14·6	71·9	75·2	10·8	12·1	*6*	M	2	112
3·5	2·9	M	94·1	18·3	13·8	75·4	18·1	13·8	—	71·2	11·8	10·7	*6*	S	5	90·7
3·6	1·1	R	84·8	18	13·7	76·1	18	13·7	—	73·5	9·1	10·4	5	M	1	114·3
4	2·7	N	79	19·4	14·9	—	19·7	14·7	74·6	78·8	11·1	11·5	5	M	3	103·6
3·8	2·9	M	80·9	19	14·1	74·2	18·9	14·3	—	73·1	9·7	11	5	S	3	113·4
3	2·5	M	97·1	18·9	13·6	—	19·1	13·4	70·1	74·9	10·1	10·8	5	M	1	106·9
4	2·4	*N*	69·5	19·6	14·4	73·5	19·4	14·3	—	72·6	10	12·1	5	M	1	121
3·5	2·1	N	90·2	19	15	78·9	19	14·8	—	70·6	9·1	11·3	5	P	1	124·2
2·7	?	M	82·1	19·3	14·3	74·1	19	14·3	—	75·4	11·2	11·7	5/6	M	2	104·5
3·2	2·5	M	75·4	20·6	15	72·6	20·4	15	—	82·1	10·1	12·2	5	M	2	120·8
4·7	2·8	M	86·2	18·8	14·7	78·2	18·8	14·5	—	74·9	9·6	11·5	5	M	1	119·8
3·7	2·5	M	86·6	19·9	15·5	—	20·4	15·3	75	72·9	10·6	12·5	5	S	4	117·9
3·8	2·5	M	75·2	20	14	—	20·1	14·1	70·1	80·1	10·2	12·5	5	M	2	122·5
3	2·2	M	94·1	19·9	14·5	72·9	19·9	14·5	—	76·2	10	12·1	5	P	1	121
2·5	3·3	M	78·2	18·9	14·1	79·9	18·8	14·3	—	81·4	10·4	10·7[a]	5	M	1	102·9

colour, identification). 117. Bones light colour, identification. 118. Sacral canal open 1st and 2nd sac. vert. behind. 119. Exostoses, 1st sacral vert. left. 120. 3rd, 4th and 5th sacral segments broken on right. 122. Four sac. vert., sacrum complete. 123. False promontory. Slight reverse curve of sacrum. 124. Six sac. vert., all sacral, double promontory. 125. Six sac. vert., all sacral, false promontory. 127. Double promontory, "Male type" of pelvis. 130. Small outlet. 132. Transitional vertebra, right sacralized, left free, false promontory.

Table of Pelvic Measurements of American Indian Squaws (continued).

No.	U. S. N. M. No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Inlet: Oblique Diameter	Inlet: Obstetric Conjugate	Inlet: Difference	Inlet: Index	Pelvic: Antero-posterior Diameter	Fetal Head	Inter-tubers	Posterior Sagittal	Pubes to Perpendicular
94	246,968	New Mexico	27·5	23·8	14·4	13·3	10·8	2·5	75	12·6	11	10	8	5·75
95	225,215	Pah Ute	25·5	22·2	12·7	12·5	10·5	2	82·7	13·5	14	11·1	11·7	2·8
96	204,254	Peru	27·3	21·5	13·8	12·2	9·25	2·9	67	11·3	10·25	9	7·5	5·3
97	225,213	”	26·5	22·25	12·5	10·4	8[a]	2·4	64	11·9	10	8·6	8·1	5·5
98	225,250	”	25	20·5	12·9	12	9·5	2·5	73·6	12·5	11	10·3	7·7[a]	6
99	227,006	”	24·6	20	12	12·3	10·8	1·5	90	11·1	10	9·6	6·3	5·7
100	227,007	”	27	24·5	13·5	14	11·6	2·4	86	11·75	12	11·75	7·4	5·7
101	227,011	”	25·3	23·3	12·3	12·1	10·5	1·6	85·4	12·5[a]	11[a]	9·5[a]	8·4[a]	4·1[a]
102	227,014	”	26	22	13·2	12·8	11·25	1·6	85·2	11·3[a]	11[a]	10	7[a]	5·5
103	227,024	”	27·3	26·5	14·3	13·1	11·35	1·8	79·4	13·6	12·5	11·5	8·8	6·1
104	227,432	”	24·5	22·3	12	11·4	9[a]	2·4	75	11·5	9·5	8·1	7	5·6
105	242,513	”	25	21·5	11·5	11·2	9·1	2·1	79·1	11·5[a]	11·5[a]	11	7·2[a]	5·3
106	242,568	”	24·5	22	11·8	11·6	9·4	2·2	79·7	11·1	9	8·3	6·3	6·4
107	225,217	Sioux	27·1	23	13·1	13·2	11·6	1·6	88·5	11·4	12	10·7	8·5	3·9
108	225,218	”	28·5	25	13·9	13·3	10·1	3·2	72·7	11·8	11	9·7	8	5·1
109	225,221	”	27·8	24·1	12·1	13·3	11·1	2·2	91·7	11	11	9·8	7·6	4·7
110	225,242	”	28	25·5	12·6	13·2	11·2	2	88·8	12·4	11·5	9·7	8·4	5·1
111	225,261	”	27	23·5	13	12·3	10	2·3	76·9	10·8	11	9·5	7·2[a]	4·1
112	225,415	”	29·1	25	14·1	14·4	12	2·4	85·1	13·2	11·25	10·6	7·6	6·5
113	225,416	”	28	24·1	12·75	11·8	10·15	1·7	79·6	12·5	10·5	9·4	7·75	6·4
114	169,672	Virginia	28[a]	24·5[a]	13·2[a]	11·8	10·2[a]	1·6	77·3	11·1	11	9·9	7·2	4·5
115	98,470	Wisconsin	25·5	22·6	13·25	?	10·85	?	81·9	10·8	10·3	9·25	7·15	5·3
	A. M. N. H., N.Y.													
116	H, 3658	Pueblo, N.M.	27	23·25	13	13·3	11·2	2·1	86·2	12·7	12·5	11	9·2	5
117	”	”	24·5	21·8	12·5	12·2	11	1·2	88	10·6	10·5	10·2	6·5	5·5
118	”	”	25·2	23	13·4	12·5	10·7	1·8	79·9	11·8	12·5	11·1	7·6	5·3
119	”	”	24·9	23[a]	13·1	11·4	9·5	1·9	72·5	12	10	9	7·4	5·9
120	”	”	27·3	25·4	13·3	12·3	10·4	1·9	77·4	12	10·5	8·5	7·5	5·2
121	H, 102	South Utah	23·8	19·3	12·4	11	9	2	72·6	11·6	12·5	11	9·4	4
122	H, 191	”	22	20	12	11·8	10·4	1·4	86·7	11·3	11	10·4	7·7	4·8
123	H, 271	”	26·5	23	13·1	11·8	10·1	1·7	77·1	11·8	10	9·6	7·2	5·8
124	H, 307	”	27	23·4	14	12·1	10·6	1·5	75·7	11·6	10·5	9·7	7·1	6·2
125	H, 321	”	25·7	22·9	13·1	10·2	9	1·2	68·7	10·3	10·5	9·7	6·8	4·7
126	99/2393	Mexico	24·5	22	12	11·6	9·9	1·7	82·5	11·2	10·5	9·5	7·1	5·7
127	99/2230	”	25	20·5	12·9	13·3	11·8	1·5	91·5	11·4	10	9	7	5·5
128	99/2165	”	26	23·7	12·5	11·5	10	1·5	80	11	10·5	8·9	7·5	5
129	99/305	”	25·6	22	13·2	11·2	9·6	1·6	72·7	10·3	10	9·2	6·6	4·6
130	99/2506	”	27	24·1	12·1	13·1	11·5	1·6	95	11·8	9·5	8·2	7·3	6
131	99/2426	”	26·9	24·6	13·1	12·1	10·5	1·6	80·1	10·2	9·5	9·2	6·1	5·3
132	99/2542	”	25·6	22	12·8	13·2	11·9	1·3	93	11·2	10	9·2	6·7	5·4
133	99/2330	”	25·1	23·5	13	14·9	13	1·9	100	12·6	11	9·5	8·5	5·3
134	99/2325	”	25·1	21·4	13	12·8	11·3	1·5	86·9	11·6	11	10	7·7	5·2
135	99/2539	”	28	24	13·5	12·8	11	1·8	81·5	12·7	11	9·1	8·1	5·8
136	99/2329	”	25·1	23·1	13·1	13·7	11·6	2·1	88·5	12·5	11	9·4	8·3	4·5
137	99/2544	”	26·1	21	12·5	12·9	11	1·9	88	10·3	10·5	9·7	8	4·4
138	99/2486	”	23·2	22	12·6	11·9	10	1·9	79·4	11·5	10	9	7	4·6

No.	Outlet: Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Innominate Bone: Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index	Pelvic Index	Sacrum: Height	Breadth	Number Segments	Curve	Segment begins at	Index
94	3·25	3	M	79·1	20·8	15·35	73·8	20·7	15·25	-	72	11·9	11·9	6	P	2	100
95	3·75	3·1	R	82·2	19·8	14·8	—	19·9	14·9	74·9	77·6	7·5	11·6	5	P	1	154·7
96	4·25	3·15	N	79·6	20·7	15·2	—	20·7	15·3	73·9	76·2	10·7	12·4	5	M	3	115·9
97	3·25	2·4	N	72·3	18·4	14·5	73·4	18·3	14·2	—	69·4	8·85	10·7	5	M	2	120·9
98	3	2·2	M	82·4	19·8	15	—	20	15·1	75·5	80	10·5[a]	11·9	6	S	2	113·3
99	3	3·3	M	86·5	18·7	14·4	77	18·4	14·4	—	76	9·7	10·5	6	M	1	108·2
100	3·25	3·2	M	100	20·1	14·5	—	20·2	14·8[a]	73·3	74·8	10·6	11·7	5	M	3	110·4
101	3·25[a]	3[a]	M	76	18·2	14	76·9	18·2	14	—	70·2	9[a]	11·5	5	M	1	127·8
102	3·25	2·8	M	88·5	18·7	14·6	—	18·7	14·7	78·6	71·9	9·4[a]	11·5	6	S	1	111·7
103	3·25	3·5	R	87·4	20·2	15·6	—	20·4	15·1	74	74·7	10	11·7	5	M	1	117
104	3·1	4	N	70·4	18·6	13·75	73·9	18·3	13·75	—	75·9	9	11·2	6	M	2	124·4
105	3·25	3	R	95·7	17·4	13·25	—	17·4[a]	13·6	78·1	69·6	8·7[a]	10	5	M	1	114·9
106	2·75	3	M	74·8	18	14·1	78·3	18	14·1	—	73·5	9·8	10·5	5/6	M	2	107·7
107	3·75	3·1	R	93·9	20·4	16·1	—	20·5	16·1	78·5	75·6	9·9	13·2	5	M	1	133·3
108	4·75	3·4	N	82·3	21·3	16·6	—	21·4	15·8	73·8	75·1	9·6	11·3	5	P	1	117·7
109	3·5	3·3	N	89·1	20·1	15·6	77·6	21·1	15·5	—	75·9	10·8	11·5	5/6	M	3	106·6
110	4	3·2	N	78·2	20·2	15·5	—	20·5	15·3	74·6	73·2	10·5	12·1	5	S	1	115·2
111	3·25	2·9	M	87·9	20	15·1	75·5	19·9	15·1	—	74·1	9·8	11·5	5	M	1	117·3
112	3	3·7	N	80·3	21·5	16·9	78·6	21·5	16·8	—	73·8	12	12·15	5	M	2	101·2
113	3·4	3·8	N	75·2	20	16	80	19·4	15·85	—	71·4	8·1	12	5	P	1	148·1
114	2·75	3	M	89·2	18·8	15	—	19·1	15[a]	78·5	68·2	9·9	11·9	5	M	3	120·2
115	2·8	3·25	N	85·6	20	14·7	73·5	19·7	14·4	—	78·4	10·5	11·8	5	M	2	112·4
116	4·3	3·3	M	86·6	20·1	15·4	—	20·1	15·5	77·1	74·4	10·1	11·4	5	S	3	112·9
117	3·7	2·5	M	96·2	18·5	14·1	—	18·5	14·3	77·3	75·5	9·7	11·5	5	M	2	117·5
118	3·3	2	R	94·1	18·8	15·4	—	18·8	15·6	77·7	74·6	10·4	11·5	5	M	3	110·6
119	3·8	1·9	M	75	18·9	14·2[a]	—	19	14·2[a]	74·7	76·3	10[a]	11·9	5	M	1	119
120	3	1·6	M	70·8	18·2	14·7	—	19·1	15·2	79·6	70	11·3	11·7	5	S	2	103·5
121	4	2·6	R	94·8	18·3	13·3	72·7	18·2	13·3[a]	—	76·9	9·2	11	5	M	3	119·6
122	4	1·2	M	92	19·6	14·2	—	19·8	14·2	71·7	90	8·2	10·4	4	M	1	126·8
123	3·5	2·1	R	84·8	19·2	14·8	77·1	19·2	14·8	—	72·5	11·3	11·3	5	Reverse	5 (?)	100
124	4·1	2·2	M	83·6	20	14·3	—	20·3	14·6	71·9	75·2	10·8	12·1	6	M	2	112
125	3·5	2·9	M	94·1	18·3	13·8	75·4	18·1	13·8		71·2	11·8	10·7	6	S	5	90·7
126	3·6	1·1	R	84·8	18	13·7	76·1	18	13·7	—	73·5	9·1	10·4	5	M	1	114·3
127	4	2·7	N	79	19·4	14·9	—	19·7	14·7	74·6	78·8	11·1	11·5	5	M	3	103·6
128	3·8	2·9	M	80·9	19	14·1	74·2	18·9	14·3	—	73·1	9·7	11	5	S	3	113·4
129	3	2·5	M	97·1	18·9	13·6	—	19·1	13·4	70·1	74·9	10·1	10·8	5	M	1	106·9
130	4	2·4	N	69·5	19·6	14·4	73·5	19·4	14·3	—	72·6	10	12·1	5	M	1	121
131	3·5	2·1	N	90·2	19	15	78·9	19	14·8		70·6	9·1	11·3	5	P	1	124·2
132	2·7	1	M	82·1	19·3	14·3	74·1	19	14·3	—	75·4	11·2	11·7	5/6	M	2	104·5
133	3·2	2·5	M	75·4	20·6	15	72·6	20·4	15	—	82·1	10·1	12·2	5	M	2	120·8
134	4·7	2·8	M	86·2	18·8	14·7	78·2	18·8	14·5	—	74·9	9·6	11·5	5	M	1	119·8
135	3·7	2·5	M	86·6	19·9	15·5	—	20·4	15·3	75	72·9	10·6	12·5	5	S	4	117·9
136	3·8	2·5	M	75·2	20	14	—	20·1	14·1	70·1	80·1	10·2	12·5	5	M	2	122·5
137	3	2·2	M	94·1	19·9	14·5	72·9	19·9	14·5		76·2	10	12·1	5	P	1	121
138	2·5	3·3	M	78·2	18·9	14·1	79·9	18·8	14·3	—	81·4	10·4	10·7[a]	5	M	1	102·9

94. Six sac. vert. (Plate IV, 4). False promontory. Spinal canal of sacrum open. 95. Shallow pelvis, wide outlet (Plate III, 3). 98. Six sac. vert., one from coccyx (?). 99. Six sac. vert., all sacral in character (Plate IV, 4). 100. False promontory (Plate VI, 6). 102. Six sac. vert., one from coccyx (?). 104. Six sac. vert., one from coccyx (?). 105. Last sacral segment missing. 106. Transitional vertebra. No false promontory. Small pelvic cavity (Plate IV, 4, Plate VI, 7). 107. Large pelvis (Plate I, 1). 108. Large pelvis. 109. Transitional vertebra, left=6, right=5 segments (Plate IV, 4). 111. Coccyx ossified to sacrum (not included). 116. Wide outlet (bones dark colour, identification). 117. Bones light colour, identification. 118. Sacral canal open 1st and 2nd sac. vert. behind. 119. Exostoses, 1st sacral vert. left. 120. 3rd, 4th and 5th sacral segments broken on right. 122. Four sac. vert., sacrum complete. 123. False promontory. Slight reverse curve of sacrum. 124. Six sac. vert., all sacral, double promontory. 125. Six sac. vert., all sacral, false promontory. 127. Double promontory, "Male type" of pelvis. 130. Small outlet. 132. Transitional vertebra, right sacralized, left free, false promontory.

Table of Pelvic Measurements of

No.	A. M. N. H., N. Y., No.	Tribe or Locality	Inter-crests cms.	Inter-spines	INLET: Transverse	Oblique Diameter	Obstetric Conjugate	Difference	Index	PELVIC: Antero-posterior Diameter	Fetal Head	Inter-tubers	Posterior Sagittal	Pubes to Perpendicular
139	H, 15075	Apache (Arizona)	27	21·4	13·3	14·3	11·5	2·8	86·5	12·8	12·5	10·5	8·9	5·2
140	H, 15098	Pueblo	26·5	23·5[a]	14	11·3	9·5	1·8	67·8	12	11·5	10	8·5	5·5
141	H, 15124	Arizona	27	23·7	13·1	11·1	9·4	1·7	71·7	10·8	11	10·2	7	5
142	H, 15094	"	24	21	12·1	12·4	10·5	1·9	86·8	11·9	11	9·9	7·9	4·6
143	H, 15154	"	25·5	22	14	10·8	9·2	1·6	65·7	12[a]	11[a]	10·2	7·3	5·3
144	H, 15124	"	25	22·2	13	10·6	9·2	1·4	70·7	11·4	10	10	6·6	6·3
145	H, 16035	S.E. Utah	24	19	11·6	11·8	10·9	0·9	93·9	10	*9*	8·4	6·6	4·5
146	99/106	Eskimo	24·5	21·8	12·1	9·4	*7·5*	1·9	61·2	11·6	11	8·8	8	4·1
147	99/3743	Brit. Columbia	28	24·1[a]	*14*	*13·6*	12·3	1·3	87·9	13·4	11·5	11·2	6·6	*7·5*
148	99/3737	"	25·4	23	13·5	13·5	11·8	1·7	87·4	11·4	11	10·5	7·5	5
149	99/1614	Queen Charlotte Island	24	20·5	11·7	12·3	10·8	1·5	92·3	10·6	10·5	10	7·1	5
150	99/3756	Brit. Columbia	27·6	24·4	13	12·4	10·5	1·9	80·8	13	11·5	10·1	8·5	6
151	99/1720	Fort Rupert, B.C.	23·7	19·5	13	11·3	9	2·3	86·9	12	12	10·3	8·2	5·2
152	99/1731	"	25·8	24	13·8	11·8	10·5	1·3	76·1	10·8	10·5	10·3	6·3	6
153	99/1727	"	24	22	13·1	13·4	11	2·4	84	12·8	12	9·9	*9*	5
154	99/1741	"	24·8	22	13·4	11·8	10·1	1·7	75·4	12·5	11·5	10·1	7·5	5·5
155	99/1670	Nimpkish, B.C.	25·5	23	13	14·3	12·9	1·4	99·2	10·9	11	10·1	7	4·7
156	99/1674	"	24·5	22·5	13·4	13·1	12	1·1	86·9	13	11·5	10 3	8	6
157	99/1676	"	26[a]	24·2[a]	14	12·4	11	1·4	78·6	*14·2*	*13*	11·2	9·5	6·1
	U. S. N. M.													
158	225,407	Sioux	25·8	22	11·8	12·4	10·8	1·6	91·5	11·5[a]	9·5[a]	8·5[a]	6·5[a]	5·3
159	225,412	"	26	22	13·3	11·7	9·5	2·2	71·4	10·4	10	9·2	6·75	4·6
160	225,414	"	26·5	24	13·5	13·1	11·8	1·3	87·4	11·3	10·5	9·3	7·9	4·8
161	225,408	Dakota	26·8	22	13·1	13·3	11	2·3	84	10·7	11	10·1	7·8	4·4
162	225,409	"	25·3	21·5	13·1	12·7	10·5	2·2	80·1	11·7	12·5	11·4	8·8	3·75
163	225,406	Cheyenne	25·5	21	13	12·9	11·3	1·6	86·9	12·1	11	10·3	7·4	6·1
164	261,810	Arkansas	27	25	14·2	10·8	9	1·8	64·1	11·8	*9*	8·4	6·5	5·8
165	261,824	"	28·2	23·5[a]	14·5	12·7	11	1·7	75·9	13·2	11	10·5	7·5	7
166	*1, C	Argentina	26	22	12·8	13·5	12	1·5	93·8	12·7	10·5	9·7	7	6·6
167	262,577	Arkansas	26·75	23	13·5	12	10	2	74·1	11·8	11	9·7	7·6	4·7
168	262,570	"	27	23·7	12·5	12·7	11	1·7	88	11·3	10·5	9	7·5	4·5
169	264,488	Peru	26·1	23·5	13	13·5	11·5	2	88·4	10·6	10·5	10·2	6·8	4·8
170	264,489	"	26·6	23·9	12·5	12·3	10·3	2	82·4	12·3	11·5	10·8	7·3	6·6
171	*18	Arizona	25	23	11·9	10·9	9·1	1·8	76·4	10·8	*9*	8·8	6·3	5·6
172	*70	"	23·2	19·1	12·8	10·8	9	1·8	70·3	11	10	8·3	7·5	4·3
173	*83	"	25	23·5	11·8	11·1[a]	9	2·1	76·2	11·6	10·5	9·4	7·3	5·3
174	*161	"	24·4	19·1	13	10·4	*8·3*	2·1	63·8	10	11	10	6	4·7
175	*2	"	27·1	24·5	12·8	11·2	9·1	2·1	71·1	12·4	10·5	*8*	*8·5*	5·2
176	*3	"	24·5	22	12·5	11·3	9·5	1·8	76	10·8[a]	11	9·4	7·5[a]	4·6
177	*104	"	26	23·5	12·8	11·4	9·2	2·2	71·9	11·4	9·5	8·6	7·3[a]	5·2
178	*112	"	25·5	22·3	12·9	11·7	9·9	1·8	76·7	11·2	10·5	9	7·5	5
179	*82	"	25·3	23·5	13·5	11·8	9·7	2·1	71·9	11·8	10·5	9·7	7·2	6·2
180	*7	"	25	22·5	12·9	11·5	9	2·5	69·7	10·1	10·5	9·7	7	4·1
181	*71	"	23·8[a]	20·9	12	11·4	9·1	2·3	75·8	10·3	11	9·9	7·7	5·8
182	*5	"	25	23	12·1	11·1	9	2·1	74·4	10·8	11	9·7	8·3	4·4
183	*Cave 3	"	27·8	24	13·9	11·5	9·2	2·3	67	10·5	10·5	10	7·2	4·8
184	*152	"	25·9	22	13·5	12·5	10·5	2	77·8	11·1	11	10·2	7·7	4·7
185	*52	"	24·5	22·2	13	10·1	*8·4*	1·7	64·6	10·4	10·5	9·4	7·5	4·1

139. Transitional vertebra, left sacral, right free. Large pelvis. 145. Small pelvis. 146. Flat pelvis. 147. Large pelvis. Last lumbar articulates with right wing sacr. 148. Six sac. vert., 1st tends toward lumbar type. False promontory. 156. Six sac. vert., last tends toward coccyx. 157. Transitional vertebra, right free, left sacralized. 158. Six sac. vert., one from lumbar. False

* Original number. Specimen not yet permanently catalogued.

American Indian Squaws (continued).

Outlet				Innominate Bone						Pelvic Index	Sacrum					
Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index		Height	Breadth	Number Segments	Curve	Segment begins at	Index
4·2	3	M	82	20·4	*16*	—	20·6	*15·3*	74·2	76·3	11·6	12·6	4/5	M	3	108·6
4·5	2·7	M	83·3	19·5	14·8	75·9	19·5	14·8[a]	—	73·6	8·2	11·9	5	M	1	145·1
3·3	3	R	94·4	19·2	15·2	79·1	19·2[a]	14·5	—	71·1	11	11·8	5	M	3	107·3
3·4	2·1	R	83·2	18	13·3	—	18·2	13·2	72·5	75	10·2	11·2	5	S	4	109·8
3	3	R	85	18·4	13·8[a]	75	18·3	13·8	—	72·5	10·5[a]	11·4	5	S	1	108·6
3·3	2·6	M	87·7	18·2	13·8	—	18·3	13·7	74·8	73·2	10·2	11·1	5	S	1	108·8
3	2·1	M	84	17·7	13·4	—	17·9	13·6	76	74·6	8·2	10·6	5	M	1	129·3
3·5	3	M	75·9	17·8	13·5	—	18	13·3	73·9	73·5	9·3	11·6	5	S	3	124·7
2·8	3·2	M	83·6	20·5	16·2	—	20·7	16·2[a]	77·8	73·9	10·1	12·3	5	M	3	121·7
3·3	2	R	92·1	20	14·8	—	20·4	14·6	70	80·3	12·6	12·7	*6*	M	2	100·8
2·8	2·1	R	94·3	18	13·5	—	18·6	13·2	70·9	77·5	11·7	11·2	*6*	M	3	95·7
3·8	2·5	M	77·7	21·9	16·6	75·8	21·6	16·1	—	79·3	10	11·7	5/6	M	2	117
3·3	1·8	R	85·8	18·6	14·5	—	18·7	14·8	—	78·9	9·4	11·4	5	M	1	121·3
3	2·5	R	95·4	20·5	15·4	75·1	20·3	15·4	—	79·4	9·2	12·2	5	P	1	132·6
3·3	2·4	R	77·3	20	14·1	70	19·7	15	—	83·3	11·1	11·9	5	M	3	107·2
3·2	3·1	R	80·1	19·6	14·6	74·5	19·5	14·7	—	79	8·5	12·3	5	P	2	144·7
3·3	2·3	R	92·7	19·8	14·7	—	19·9	15	75·4	78	9·6	10·3	5	S	4	107·3
3·7	2·9	R	79·2	20	14·6[a]	70·3	19·8	14·6	—	81·6	7·6	11·9	5	M	2	151·1
3·3	2·7	R	78·9	20·6[a]	15·5[a]	—	20·6	15·5	75·2	79·2	10·4	12·3	5	M	1	118·3
3	2·6	N	73·9	19	14·7	—	19·2	14·5	75·5	74·4	12[a]	11·8	*6*	M	2	98·3
3·25	2·9	M	88·4	19·6	15	76·5	19·6	15	—	75·4	10·1	11·3	5	M	1	111·8
3·5	3·3	M	82·3	20·5	16·2	—	20·6	16	77·7	77·7	10·9	10·7	5	M	3	98·2
3·75	2·7	R	94·4	20·3	16·4	—	20·4	15·7	77	76·1	10·2	12	5	S	1	117·6
3	3·4	R	97·4	19·4	14·8	—	19·7	15·1	75·6	77·8	10·6	11·4	5	M	2	107·6
3·5	3·3	M	85·1	20	15·6	73	19·9	15·4	—	78·4	9·1	11·2	5	P	1	123·1
2·5	2·9	N	71·2	20·2	14·3	70·8	19·8	14·3[a]	—	74·8	12·1	12·1	*6*	M	1	100
3·5	3·4	M	79·5	21·4	15·5[a]	—	21·6	15·5	71·8	76·6	11·7	12·4	5	M	3	106
3	2·8	N	76·4	20	15·3	—	20·3	15·1	78·3	78·3	11·9	11·8	*6*	M	3	99·1
3·5	2·8	R	82·2	19	15·1	79·5	19	14·9	—	71	11·6	12·5	*6*	M	1	107·8
3·5	2·9	N	79·6	20	14·6	73	19·9	14·7	—	74·1	10·8	12·7	5	M	2	117·6
3·25	2·4	R	96·2	19	14·2	74·7	18·8	14·8	—	72·4	9·9	11·1	*6*	M	1	112·1
3·75	2·8	M	87·8	20	14·9	—	20	15	75	75·2	9·4	11·2	5	M	2	119·1
3·3	2·5	N	81·5	17·7	13·2	74·6	17·7	13·2[a]	—	70·8	8·5	10·1	5	M	3	118·8
2·5	3·2	M	75·4	17·4	13·6	72·4	17·3	13·6[a]	—	75	9·6	11	5	S	4	114·6
3·2	2·4	N	81·0	17·1	13·4	78·4	17·2	13·4	—	68·8	10·2	10·8	5	S	4	105·9
3·3	2·3	R	100	17·2	13	75·6	17·2[a]	13[a]	—	70·5	9·2	10·4	5	M	2	113
3·6	3·1	*N*	64·5	19	13·5	71·1	18·7	14·5	—	70·1	8·4	11	5	M	2	131
3·6	3	M	87	18	13·7	—	18·1	13·7[a]	75·7	73·9	9·2[a]	11·1	5	M	2	120·7
3·3	3·1	N	75·4	18·9	14·6	77·2	18·9[a]	14·6[a]	—	72·7	8·5[a]	12	5	M	2	141·1
3·6	2·3	N	80·4	18·4	13·3	—	18·7	13·2	75·9	73·3	10·1	11·1	5	M	1	109·9
3·6	3	N	82·2	19·6	14	—	19·8	14·3	72·2	74·3	10·5	11·6	5	M	2	110·5
3·2	2·9	R	96	18·1	13[a]	71·8	17·8	13[a]	—	72·4	9·3	11	5	M	2	118·3
3·4	2[a]	R	96·1	17·7	13	—	17·7	13·2	74·6	74·3	9·8	10·7	5	S	3	109·2
4	2·9	R	89·8	18·3[a]	13·7	—	18·4	13·9	75·5	73·6	9	10·2[a]	5	M	1	113·3
4	3	M	95·2	19·5	14·2	—	19·5	14·3	73·3	70·1	9·7	12	5	S	4	123·7
3·7	2·6	R	91·9	20	14·5	72·5	19·9	14·3	—	77·2	9·3	12·1	5	M	1	130·1
3·5	2·2	R	90·8	17·8	13·7	78·1	17·5	13·6	—	72·7	9	10·7	5	M	3	118·9

promontory. 160. False promontory, 1st sacral has lumbar characteristics. 162. Wide outlet. Coccyx ossified to sacrum (not included). 164. Six sac. vert., one from coccyx (?). Small outlet. 166. Six sac. vert., all sacral. Double promontory, rounded. 167. Six sac. vert., all sacral. 169. Six sac. vert., all sacral in character. 174. Flat pelvis. 175. "Male arch." 181. Youth. Sacral bodies not quite united.

Table of Pelvic Measurements of American Indian Squaws (continued).

No.	A. M. N. H., N. Y., No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Oblique Diameter	Obstetric Conjugate	Difference	Index	Pelvic: Antero-posterior Diameter	Fetal Head	Inter-tubers	Posterior Sagittal	Pubes to Perpendicular
139	H, 15075	Apache (Arizona)	27	21·4	13·3	14·3	11·5	2·8	86·5	12·8	12·5	10·5	8·9	5·2
140	H, 15098	Pueblo	26·5	23·5[a]	14	11·3	9·5	1·8	67·8	12	11·5	10	8·5	5·5
141	H, 15124	Arizona	27	23·7	13·1	11·1	9·4	1·7	71·7	10·8	11	10·2	7	5
142	H, 15094	,,	24	21	12·1	12·4	10·5	1·9	86·8	11·9	11	9·9	7·9	4·6
143	H, 15154	,,	25·5	22	14	10·8	9·2	1·6	65·7	12[a]	11[a]	10·2	7·3	5·3
144	H, 15121	,,	25	22·2	13	10·6	9·2	1·4	70·7	11·4	10	10	6·6	6·3
145	H, 16035	S.E. Utah	24	19	11·6	11·8	10·9	0·9	93·9	10	*9*	8·4	6·6	4·5
146	99 406	Eskimo	24·5	21·8	12·1	9·4	*7·5*	1·9	61·2	11·6	11	8·8	8	4·1
147	99 3743	Brit. Columbia	28	24·1[a]	*14*	*13·6*	12·3	1·3	87·9	13·4	11·5	11·2	6·6	*7·5*
148	99 3737	,,	25·4	23	13·5	13·5	11·8	1·7	87·4	11·4	11	10·5	7·5	5
149	99 1614	Queen Charlotte Island	24	20·5	11·7	12·3	10·8	1·5	92·3	10·6	10·5	10	7·1	5
150	99 3756	Brit. Columbia	27·6	24·4	13	12·4	10·5	1·9	80·8	13	11·5	10·1	8·5	6
151	99/1720	Fort Rupert, B.C.	23·7	19·5	13	11·3	9	2·3	86·9	12	12	10·3	8·2	5·2
152	99/1731	,,	25·8	24	13·8	11·8	10·5	1·3	76·1	10·8	10·5	10·3	6·3	6
153	99/1727	,,	24	22	13·1	13·4	11	2·4	84	12·8	12	9·9	9	5
154	99/1741	,,	24·8	22	13·4	11·8	10·1	1·7	75·4	12·5	11·5	10·1	7·5	5·5
155	99/1670	Nimpkish, B.C.	25·5	23	13	14·3	12·9	1·4	99·2	10·9	11	10·1	7	4·7
156	99/1674	,,	24·5	22·5	13·4	13·1	12	1·1	86·9	13	11·5	10 3	8	6
157	99/1676	,,	26[a]	24·2[a]	14	12·4	11	1·4	78·6	*14·2*	*13*	11·2	9·5	6·1
	U. S. N. M.													
158	225,407	Sioux	25·8	22	11·8	12·4	10·8	1·6	91·5	11·5[a]	9·5[a]	8·5[a]	6·5[a]	5·3
159	225,412	,,	26	22	13·3	11·7	9·5	2·2	71·4	10·4	10	9·2	6·75	4·6
160	225,414	,,	26·5	24	13·5	13·1	11·8	1·3	87·4	11·3	10·5	9·3	7·9	4·8
161	225,408	Dakota	26·8	22	13·1	13·3	11	2·3	84	10·7	11	10·1	7·8	4·4
162	225,409	,,	25·3	21·5	13·1	12·7	10·5	2·2	80·1	11·7	12·5	11·4	8·8	3·75
163	225,406	Cheyenne	25·5	21	13	12·9	11·3	1·6	86·9	12·1	11	10·3	7·4	6·1
164	261,8[illegible]0	Arkansas	27	25	14·2	10·8	9	1·8	64·1	11·8	*9*	8·4	6·5	5·8
165	261,824	,,	28·2	23·5[a]	14·5	12·7	11	1·7	75·9	13·2	11	10·5	7·5	7
166	*1, C	Argentina	26	22	12·8	13·5	12	1·5	93·8	12·7	10·5	9·7	7	6·6
167	262,577	Arkansas	26·75	23	13·5	12	10	2	74·1	11·8	11	9·7	7·6	4·7
168	262,570	,,	27	23·7	12·5	12·7	11	1·7	88	11·3	10·5	9	7·5	4·5
169	264,488	Peru	26·1	23·5	13	13·5	11·5	2	88·4	10·6	10·5	10·2	6·8	4·8
170	264,489	,,	26·6	23 9	12·5	12·3	10·3	2	82·4	12·3	11·5	10·8	7·3	6·6
171	*18	Arizona	25	23	11·9	10·9	9·1	1·8	76·4	10·8	*9*	8·8	6·3	5·6
172	*70	,,	23·2	19·1	12·8	10·8	9	1·8	70·3	11	10	8·3	7·5	4·3
173	*83	,,	25	23·5	11·8	11·1[a]	9	2·1	76·2	11·6	10·5	9·4	7·3	5·3
174	*161	,,	24·4	19·1	13	10·4	*8·3*	2·1	63·8	10	11	10	6	4·7
175	*2	,,	27·1	24·5	12·8	11·2	9·1	2·1	71·1	12·4	10·5	*8*	*8·5*	5·2
176	*3	,,	24·5	22	12·5	11·3	9·5	1·8	76	10·8[a]	11	9·4	7·5[a]	4·6
177	*104	,,	26	23·5	12·8	11·4	9·2	2·2	71·9	11·4	9·5	8·6	7·3[a]	5·2
178	*112	,,	25·5	22·3	12·9	11·7	9·9	1·8	76·7	11·2	10·5	9	7·5	5
179	*82	,,	25·3	23·5	13·5	11·8	9·7	2·1	71·9	11·8	10·5	9·7	7·2	6·2
180	*7	,,	25	22·5	12·9	11·5	9	2·5	69·7	10·1	10·5	9·7	7	4·1
181	*71	,,	23·8[a]	20·9	12	11·4	9·1	2·3	75·8	10·3	11	9·9	7·7	5·8
182	*5	,,	25	23	12·1	11·1	9	2·1	74·4	10·8	11	9·7	8·3	4·4
183	*Cave 3	,,	27·8	24	13·9	11·5	9·2	2·3	67	10·5	10·5	10	7·2	4·8
184	*152	,,	25·9	22	13·5	12·5	10·5	2	77·8	11·1	11	10·2	7·7	4·7
185	*52	,,	24·5	22·2	13	10·1	*8·4*	1·7	64·6	10·4	10·5	9·4	7·5	4·1

No.	Outlet: Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Innominate Bone: Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index	Pelvic Index	Sacrum: Height	Breadth	Number Segments	Curve	Segment begins at	Index
139	4·2	3	M	82	20·4	*16*	—	20·6	*15·3*	74·2	76·3	11·6	12·6	4/5	M	3	108·6
140	4·5	2·7	M	83·3	19·5	14·8	75·9	19·5	14·8[a]	—	73·6	8·2	11·9	5	M	1	145·1
141	3·3	3	R	94·4	19·2	15·2	79·1	19·2[a]	14·5	—	71·1	11	11·8	5	M	3	107·3
142	3·4	2·1	R	83·2	18	13·3	—	18·2	13·2	72·5	75	10·2	11·2	5	S	4	109·8
143	3	3	R	85	18·4	13·8[a]	75	18·3	13·8	—	72·5	10·5[a]	11·4	5	S	1	108·6
144	3·3	2·6	M	87·7	18·2	13·8	—	18·3	13·7	74·8	73·2	10·2	11·1	5	S	1	108·8
145	3	2·1	M	84	17·7	13·4	—	17·9	13·6	76	74·6	8·2	10·6	5	M	1	129·3
146	3·5	3	M	75·9	17·8	13·5	—	18	13·3	73·9	73·5	9·3	11·6	5	S	3	124·7
147	2·8	3·2	M	83·6	20·5	16·2	—	20·7	16·2[a]	77·8	73·9	10·1	12·3	5	M	3	121·7
148	3·3	2	R	92·1	20	14·8	—	20·4	14·6	70	80·3	12·6	12·7	*6*	M	2	100·8
149	2·8	2·1	R	94·3	18	13·5	—	18·6	13·2	70·9	77·5	11·7	11·2	*6*	M	3	95·7
150	3·8	2·5	M	77·7	21·9	16·6	75·8	21·6	16·1	—	79·3	10	11·7	5/6	M	2	117
151	3·3	1·8	R	85·8	18·6	14·5	—	18·7	14·8	—	78·9	9·4	11·4	5	M	1	121·3
152	3	2·5	R	95·4	20·5	15·4	75·1	20·3	15·4	—	79·4	9·2	12·2	5	P	1	132·6
153	3·3	2·4	R	77·3	20	14·1	70	19·7	15	—	83·3	11·1	11·9	5	M	3	107·2
154	3·2	3·1	R	80·1	19·6	14·6	74·5	19·5	14·7	—	79	8·5	12·3	5	P	2	144·7
155	3·3	2·3	R	92·7	19·8	14·7	—	19·9	15	75·4	78	9·6	10·3	5	S	4	107·3
156	3·7	2·9	R	79·2	20	14·6[a]	70·3	19·8	14·6	—	81·6	7·6	11·9	5	M	2	151·1
157	3·3	2·7	R	78·9	20·6[a]	15·5[a]	—	20·6	15·5	75·2	79·2	10·4	12·3	5	M	1	118·3
158	3	2·6	N	73·9	19	14·7	—	19·2	14·5	75·5	74·4	12[a]	11·8	*6*	M	2	98·3
159	3·25	2·9	M	88·4	19·6	15	76·5	19·6	15	—	75·4	10·1	11·3	5	M	1	111·8
160	3·5	3·3	M	82·3	20·5	16·2	—	20·6	16	77·7	77·7	10·9	10·7	5	M	3	98·2
161	3·75	2·7	R	94·4	20·3	16·4	—	20·4	15·7	77	76·1	10·2	12	5	S	1	117·6
162	3	3·4	R	97·4	19·4	14·8	—	19·7	15·1	75·6	77·8	10·6	11·4	5	M	2	107·6
163	3·5	3·3	M	85·1	20	15·6	73	19·9	15·4	—	78·4	9·1	11·2	5	P	1	123·1
164	2·7	2·9	N	71·2	20·2	14·3	70·8	19·8	14·3[a]	—	74·8	12·1	12·1	*6*	M	1	100
165	3·5	3·4	M	79·5	21·4	15·5[a]	—	21·6	15·5	71·8	76·6	11·7	12·4	5	M	3	106
166	3	2·8	N	76·4	20	15·3	—	20·3	15·1	78·3	78·3	11·9	11·8	*6*	M	3	99·1
167	3·5	2·8	R	82·2	19	15·1	79·5	19	14·9	—	71	11·6	12·5	*6*	M	1	107·8
168	3·5	2·9	N	79·6	20	14·6	73	19·9	14·7	—	74·1	10·8	12·7	5	M	2	117·6
169	3·25	2·4	R	96·2	19	14·2	74·7	18·8	14·8	—	72·4	9·9	11·1	*6*	M	1	112·1
170	3·75	2·8	M	87·8	20	14·9	—	20	15	75	75·2	9·4	11·2	5	M	2	119·1
171	3·3	2·5	N	81·5	17·7	13·2	74·6	17·7	13·2[a]	—	70·8	8·5	10·1	5	M	3	118·8
172	2·5	3·2	M	75·4	17·4	13·6	72·4	17·3	13·6[a]	—	75	9·6	11	5	S	4	114·6
173	3·2	2·4	N	81·0	17·1	13·4	78·4	17·2	13·4	—	68·8	10·2	10·8	5	S	4	105·9
174	3·3	2·3	R	100	17·2	13	75·6	17·2[a]	13[a]	—	70·5	9·2	10·4	5	M	2	113
175	3·6	3·1	*N*	64·5	19	13·5	71·1	18·7	14·5	—	70·1	8·4	11	5	M	2	131
176	3·6	3	M	87	18	13·7	—	18·1	13·7[a]	75·7	73·9	9·2[a]	11·1	5	M	2	120·7
177	3·3	3·1	N	75·4	18·9	14·6	77·2	18·9[a]	14·6[a]	—	72·7	8·5[a]	12	5	M	2	141·1
178	3·6	2·3	N	80·4	18·4	13·3	—	18·7	13·2	75·9	73·3	10·1	11·1	5	M	1	109·9
179	3·6	3	N	82·2	19·6	14	—	19·8	14·3	72·2	74·3	10·5	11·6	5	M	2	110·5
180	3·2	2·9	R	96	18·1	13[a]	71·8	17·8	13[a]	—	72·4	9·3	11	5	M	2	118·3
181	3·4	2[a]	R	96·1	17·7	13	—	17·7	13·2	74·6	74·3	9·8	10·7	5	S	3	109·2
182	4	2·9	R	89·8	18·3[a]	13·7	—	18·4	13·9	75·5	73·6	9	10·2[a]	5	M	1	113·3
183	4	3	M	95·2	19·5	14·2	—	19·5	14·3	73·3	70·1	9·7	12	5	S	4	123·7
184	3·7	2·6	R	91·9	20	14·5	72·5	19·9	14·3	—	77·2	9·3	12·1	5	M	1	130·1
185	3·5	2·2	R	90·8	17·8	13·7	78·1	17·5	13·6	—	72·7	9	10·7	5	M	3	118·9

139. Transitional vertebra, left sacral, right free. Large pelvis. 145. Small pelvis. 146. Flat pelvis. 147. Large pelvis. Last lumbar articulates with right wing sacr. 148. Six sac. vert., 1st tends toward lumbar type. False promontory. 156. Six sac. vert., last tends toward coccyx. 157. Transitional vertebra, right free, left sacralized. 158. Six sac. vert., one from lumbar. False promontory. 160. False promontory, 1st sacral has lumbar characteristics. 162. Wide outlet. Coccyx ossified to sacrum (not included). 164. Six sac. vert., one from coccyx(?). Small outlet. 166. Six sac. vert., all sacral. Double promontory, rounded. 167. Six sac. vert., all sacral. 169. Six sac. vert., all sacral in character. 174. Flat pelvis. 175. "Male arch." 181. Youth. Sacral bodies not quite united.

* Original number. Specimen not yet permanently catalogued.

Table of Pelvic Measurements of

No.	U. S. N. M. No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Inlet: Oblique Diameter	Inlet: Obstetric Conjugate	Inlet: Difference	Inlet: Index	Pelvic: Antero-posterior Diameter	Pelvic: Fetal Head	Pelvic: Inter-tubers	Pelvic: Posterior Sagittal	Pelvic: Pubes to Perpendicular
186	*21	Arizona	24·7	22	12·2	11·4	9·6	1·8	78·7	11·7	11	9·7	7·8	4·7
187	*128	"	25·5	22·7	12·5	10·8	9	1·8	72	10·6	9·5	8·5	6·2	5
188	*68	"	24	21[a]	12·1	11·8	9·2	2·6	76	12	9·5	8·6	7·1	5·4
189	*73	"	27	22·4	13	11·3	9	2·3	69·2	11·5	11·5	10·7	8	4·5
190	*34	"	24·3[a]	21·4[a]	11·9	13·3	11·9	1·4	100	12	9·5	8·7[a]	7	6
191	*15	"	26	24	13	11	8·8	2·2	67·2	11·4[a]	9·5[a]	8·1	7[a]	4·6
192	*28	"	26·6	23	13	9·6	*8·1*	1·5	61·8	11·4	10·5	9·1	8	4·2
193	*1	"	26·1	23·1	13·9	10·7	9	1·7	64·7	11	10·5	9·7	7	5·8
194	*57	"	26·5	24·4	13	11·3	9·4	1·9	72·3	10·8	10·5	10·3	6·6	5·2
195	*1,b	"	24·2	20	12·3	11·5[a]	9·5[a]	2	77·2	11·7[a]	10·5	10	6·3	5·6
196	*64	"	26·2	21·5	13·2	13	10·9	2·1	82·6	11·1	10	9·2	6·5	5·6
197	*79	"	24[a]	20	12·2	11·7	10·2	1·5	83·6	10·3	9·5	8·3[a]	6·8	4·3
	A. N. H. M., N. Y.													
198	99/1677	Nimpkish, B.C.	26·9	24·2	13·4	12·6	11·2	1·4	83·6	13·2	11	9·5	8·2	6·2
199	99/1672	"	25·1	23·8	13·1	13·6	12·3	1·3	93·9	12·6	12·5	10·8	8·3	5·4
200	99/1668	"	23	19·5	*10·3*	12·8	11	1·8	106·9	11·7	11·5	10·8	7·5	5·5
201	99/1669	"	26	21·8	12·6	12·8	11·5	1·3	91·3	12	11·5	10·6	8	5·6
202	99/101	Kwakiutl	24·9	23	12·9	11·7	10	1·7	77·5	12·5	12	10	9·3	4·6
203	"2"	N. W. Coast	25	21·5	13·1	12	10·5	1·5	80·9	12·2	11	9·6	8	4·8
204	99/1619	Nanaimo V. C.	23	21·3	12	12·5	11	1·5	91·7	*9·5*	11	10	7	3·5
205	99/1620	"	26	24	13·7	12·3	10	*2·3*	73	11	11	10·5	7·5	5·4
206	99/1623	"	24	21	13·5	11·4	10	1·4	74·1	12·5	12	11	7·7	5·7
207	99/1625	"	27·8	24	14	12·6	10·5	2·1	75	13	12·5	11·5	8·5	5·5
208	99/1626	"	24·9	22·1	13·5	12	10·4	1·6	70	11·5	11	10·2	7	5·3
209	99/2637	N. Samich, B.C.	25	21·6	12·9	11·3	9·8	1·5	76	12·7	11	10·3	7	6·2
210	99/1699	"	25·9	22·8	13	13·8	12	1·8	92·3	12·9	11·5	10·5[a]	8	6·3
211	99/1520	Victoria, B.C.	26·5[a]	22[a]	13·1	12	10·4	1·6	79·4	11·2	11·5	10·9	7	4·6
212	99/2632	N. Samich	25[a]	21·5[a]	13	12·2	11	1·2	84·6	12·6	10·5[a]	10·4	7·2	6·1
213	99/2666	"	24	22·4	13	10·8	9·5	1·3	73·1	13·3	12·5	10·6	9	5
214	22,183	Wallula, Oreg.	24·9	20·8	12·9	11·4	9·9	1·5	76·7	11·1	12	10·1	8·8	3·8
	U. S. N. M.													
215	*60	Arizona	24·2	21	11·8	11·4	10	1·4	84·7	11·7	10	8·5	7	5
216	*149	"	26	23	13	12	10·3	1·7	79·2	10·5	9·5	8·8	6·5	4·8
217	*148	"	27	23·5	12·4	11	9	2	72·6	10·7	9	8·2[a]	6·3	5·5
	Average †	—	25·76	22·66	12·95	11·78	10·68	1·76	79·75	11·59	10·82	9·79	7·56	5·21
	Largest ‡	—	29·1	27	14·7	14·9	14	3·2	107·7	14·66	14	12·75	11·7	7·5
	Smallest	—	20·8	17·2	10·3	9·4	7·5	0·8	61·5	9	9	8	6	2·8

188. "Male type." 189. Sacral canal open all but 0·1 cm. in middle. 190. Narrow, somewhat asymmetrical pelvis. 191. Six sac. vert., one from coccyx? "Male type." Narrow outlet. 192. Flat pelvis. Small outlet. 194. Transitional vertebra, left sacral, right lumbar type. False promontory. 196. Six sac. vert., one from coccyx (?). 198. Six sac. vert., all sacral character.

* Original number. Specimen not yet permanently catalogued.

American Indian Squaws (continued).

Outlet				Innominate Bone						Pelvic Index	Sacrum					
Height of Perpendicular	Pubic Height	Sub-pubic Arch	Outlet Index	Height, Right	Breadth, Right	Index	Height, Left	Breadth, Left	Index		Height	Breadth	Number Segments	Curve	Segment begins at	Index
3	2·8	M	82·9	18·3	13·6	—	18·4	13·6	73·9	74·1	9·1	10·7	5	M	1	117·6
3	3	M	80·2	18·3	14	76·5	18·2	13·6	—	71·8	9·6	11·3	5	M	2	117·7
2·7	3·1	N	71·7	18·6	14·2	76·3	18·4	14[a]	—	77·5	11	11·1	5	S	3	100·9
3·5	2·8	R	93	18·3	15	82	18·3	14·6	—	67·8	9·6	11·6	5	M	1	120·8
2·8	2·2	N	72·5	19·5[a]	14[a]	71·8	19·3	14[a]	—	80·2	9·8	11·6	5	M	1	118·4
2·8	3·2	N	71·1	19·2	13·8	—	19·3	13·4	69·4	74·2	11·1[a]	11·1	6	M	2	100
3·3	3	N	79·8	18·7	14·4	—	19	14·6	76·4	71·4	10	11·1	5	S	1	111
3·5	2	M	88·2	18·1	13·8	—	18·3	13·5	73·7	70·1	9	10·9	5	M	1	101
2·8	3·8	R	95·4	20·4	15	73·5	20·3	14·8	—	77	9·7	10·8	5/6	M	2	111·3
3·4	2·5	M	85·5	18·5	14·2	76·8	18·4	14·3	—	76·4	9·4	10·4	5	M	3	110·6
3·2	2·7	M	82·9	20·2	14·6	—	20·7	14·6	70·5	79	11·4	11·4	6	M	3	100
3	3	M	80·6	19·1	14·2	74·3	19·1[a]	14·2[a]	—	79·6	9·7	11·3	5	M	1	116·5
3·2	2·8	M	72	19·3	15·2	—	19·5	15·2	77·9	72·5	10·9	12·3	6	S	4	112·8
3·3	2·8	R	85·7	20·3	15·8	—	20·5	16·1	78·5	81·7	12·3	12·3	5/6	M	3	100
3·8	1·4	M	92·3	19	14·1	74·2	18·7	14	—	82·6	11·4	12	5	S	4	105·3
3·7	2·5	M	88·3	20·1	15·8	—	20·3	15·6	76·8	78·1	9·8	12	5	M	1	122·4
3·2	2·6	R	80	18·3	14	—	18·3	14·4	78·7	73·5	9·1	11	5	S	1	120·9
3·5	3·3	N	78·7	20·2	14·8	76·5	19·7	14·6	—	80·8	8·2	11·3	5	M	1	137·8
2·8	3	R	105·3	18·3	13·6	—	18·9	13·8	73·0	82·2	10	11·3	5	M	1	113
3·4	3	M	95·5	20·8	14·6	70·2	20·4	14·5	—	80	9·8	11·5	5	M	1	117·3
3·2	2·3	R	88	17·9	14·2	79·3	17·8	14·1	—	74·6	9·3	11·6	5	S	2	124·7
3·4	2·6	R	88·5	20·7	15·8	—	20·8	15·8	76	74·8	11·5	12·7	6	M	1	110·4
3·3	1·8	R	88·7	18·5	14·7	—	18·9	14·9	78·8	75·9	10·6	12·5	5	M	2	117·9
2·5	2·5	M	81·1	18·6	14·9	80·1	18·2	14·1	—	74·4	9·7	10·3	5	S	2	106·2
3	3	R	81·4	20·1	15·8[a]	—	20·1	15·8	78·6	77·6	11·6	12	6	S	2	103·4
2·8	2·3	R	97·3	19·2[a]	14·5	—	19·2	14·5[a]	75·5	72·4	9·3	11	5	M	1	126·9
2·8	2·1	R	82·5	20·8[a]	14·8[a]	—	20·8	14·8	71·2	83·2	9·1	11·7	5/6	M	2	128·5
3·5	2·3	M	79·7	18·3	13·7	74·9	18·3[a]	13·6	—	76·2	8·2	10·7	6/5	M	3	130·5
3·1	1·8	R	91	18·7	14·2	76·5	18·6	14·1	—	73·5	9·6	11·6	5	M	1	120·8
3	2·6	N	72·6	18·1	13	—	18·6	12·8	68·8	76·4	9·7	11	5	S	3	113·4
2·4	2·5	M	83·8	19·1	14·2	—	19·2	14·2	74·5	73·8	10·8	11·7	6	M	1	108·3
3	2·4	N	76·6	19·5	14·5	74·4	19·4	14·7	—	72·2	10·2	10·9	6/5	M	3	106·8
		N-44											4 segs. = 3		1st = 90	
3·35	2·81	N-M-5	84·26	19·27	14·52	74·6	19·32	14·5	—	75·2	10	11·5	4/5 = 2	S 52	2nd = 59	115·8
4·5	4·3	M-92	108·9	22·2	16·9	84·1	22·5	16·8	—	90	13·2	13·7	5 = 169	M 146	3rd = 49	174·6
2·5	1·1	M-R-4	64	16	11·3	67·8	15·7	11·15	—	63·7	6·3	9·2	5/6 = 11	P 18	4th = 16	86
		R-71											6 = 32	reversed 1	5th = 3	

199. Transitional vertebra, free on right, one from coccyx. 200. Narrow pelvis. 207. Six sac. vert., all sacral. No false promontory. 210. Six sac. vert., all sacral. No false promontory. 212. Transitional vertebra, right free, but articulates with ilium, left sac. False promontory. 213. Transitional vertebra, left free, right sacralized. False promontory. 216. Six sac. vert., one from coccyx. 217. Transitional vertebra, right sacralized, left free. No false promontory. † False promontory = 18. ‡ Double promontory = 2.

Table of Pelvic Measurements of American Indian Squaws (continued).

No.	U. S. N. M. No.	Tribe or Locality	Inter-crests cms.	Inter-spines	Inlet: Transverse	Inlet: Oblique Diameter	Inlet: Obstetric Conjugate	Inlet: Difference	Inlet: Index	Inlet: Antero-posterior Diameter	Pelvic Outlet: Fetal Head	Pelvic Outlet: Inter-tubers	Pelvic Outlet: Posterior Sagittal	Pelvic Outlet: Pubes to Perpendicular
186	*21	Arizona	24·7	22	12·2	11·4	9·6	1·8	78·7	11·7	11	9·7	7·8	4·7
187	*128	,,	25·5	22·7	12·5	10·8	9	1·8	72	10·6	9·5	8·5	6·2	5
188	*68	,,	24	21[a]	12·1	11·8	9·2	2·6	76	12	9·5	8·6	7·1	5·4
189	*73	,,	27	22·4	13	11·3	9	2·3	69·2	11·5	11·5	10·7	8	4·5
190	*34	,,	24·3[a]	21·4[a]	11·9	13·3	11·9	1·4	100	12	9·5	8·7[a]	7	6
191	*15	,,	26	24	13	11	8·8	2·2	67·2	11·4[a]	9·5[a]	8·1	7[a]	4·6
192	*28	,,	26·6	23	13	9·6	8·7	1·5	61·8	11·4	10·5	9·1	8	4·2
193	*1	,,	26·1	23·1	13·9	10·7	9	1·7	64·7	11	10·5	9·7	7	5·8
194	*57	,,	26·5	24·4	13	11·3	9·4	1·9	72·3	10·8	10·5	10·3	6·6	5·2
195	*1,b	,,	24·2	20	12·3	11·5[a]	9·5[a]	2	77·2	11·7[a]	10·5	10	6·3	5·6
196	*64	,,	26·2	21·5	13·2	13	10·9	2·1	82·6	11·1	10	9·2	6·5	5·6
197	*79	,,	24[a]	20	12·2	11·7	10·2	1·5	83·6	10·3	9·5	8·3[a]	6·8	4·3
	A. N. H. M., N. Y.													
198	99/1677	Nimpkish, B.C.	26·9	24·2	13·4	12·6	11·2	1·4	83·6	13·2	11	9·5	8·2	6·2
199	99/1672	,,	25·1	23·8	13·1	13·6	12·3	1·3	93·9	12·6	12·5	10·8	8·3	5·4
200	99/1668	,,	23	19·5	10·3	12·8	11	1·8	106·9	11·7	11·5	10·8	7·5	5·5
201	99/1669	,,	26	21·8	12·6	12·8	11·5	1·3	91·3	12	11·5	10·6	8	5·6
202	99/101	Kwakiutl	24·9	23	12·9	11·7	10	1·7	77·5	12·5	12	10	9·3	4·6
203	*2	N. W. Coast	25	21·5	13·1	12	10·5	1·5	80·9	12·2	11	9·6	8	4·8
204	99/1619	Nanaimo V. C.	23	21·3	12	12·5	11	1·5	91·7	9·5	11	10	7	3·5
205	99/1620	,,	26	24	13·7	12·3	10	2·3	73	11	11	10·5	7·5	5·4
206	99/1623	,,	24	21	13·5	11·4	10	1·4	74·1	12·5	12	11	7·7	5·7
207	99/1625	,,	27·8	24	14	12·6	10·5	2·1	75	13	12·5	11·5	8·5	5·5
208	99/1626	,,	24·9	22·1	13·5	12	10·4	1·6	70	11·5	11	10·2	7	5·3
209	99/2637	N. Saanich, B.C.	25	21·6	12·9	11·3	9·8	1·5	76	12·7	11	10·3	7	6·2
210	99/1699	,,	25·9	22·8	13	13·8	12	1·8	92·3	12·9	11·5	10·5[a]	8	6·3
211	99/1520	Victoria, B.C.	26·5[a]	22[a]	13·1	12	10·4	1·6	79·4	11·2	11·5	10·9	7	4·6
212	99/2632	N. Saanich	25[a]	21·5[a]	13	12·2	11	1·2	84·6	12·6	10·5[a]	10·4	7·2	6·1
213	99/2666	,,	24	22·1	13	10·8	9·5	1·3	73·1	13·3	12·5	10·6	9	5
214	22,183	Wallula, Oreg.	24·9	20·8	12·9	11·4	9·9	1·5	76·7	11·1	12	10·1	8·8	3·8
	U. S. N. M.													
215	*60	Arizona	24·2	21	11·8	11·4	10	1·4	84·7	11·7	10	8·5	7	5
216	*149	,,	26	23	13	12	10·3	1·7	79·2	10·5	9·5	8·8	6·5	4·8
217	*148	,,	27	23·5	12·4	11	9	2	72·6	10·7	9	8·2[a]	6·3	5·5
	Average †	—	25·76	22·66	12·95	11·78	10·68	1·76	79·75	11·59	10·82	9·79	7·56	5·21
	Largest ‡	—	29·1	27	14·7	14·9	14	3·2	107·7	14·66	14	12·75	11·7	7·5
	Smallest	—	20·8	17·2	10·3	9·4	7·5	0·8	61·5	9	9	8	6	2·8

No.	Outlet: Height of Perpendicular	Outlet: Pubic Height	Sub-pubic Arch	Outlet Index	Innominate Bone: Height, Right	Innominate Bone: Breadth, Right	Innominate Bone: Index	Innominate Bone: Height, Left	Innominate Bone: Breadth, Left	Innominate Bone: Index	Pelvic Index	Sacrum: Height	Sacrum: Breadth	Sacrum: Number Segments	Sacrum: Curve	Sacrum: Segment begins at	Sacrum: Index
186	3	2·8	M	82·9	18·3	13·6		18·4	13·6	73·9	74·1	9·1	10·7	5	M	1	117·6
187	3	3	M	80·2	18·3	14	76·5	18·2	13·6	-	71·8	9·6	11·3	5	M	2	117·7
188	2·7	3·1	N	71·7	18·6	14·2	76·3	18·4	14[a]		77·5	11	11·1	5	S	3	100·9
189	3·5	2·8	R	93	18·3	15	82	18·3	14·6		67·8	9·6	11·6	5	M	1	120·8
190	2·8	2·2	N	72·5	19·5[a]	14[a]	71·8	19·3	14[a]		80·2	9·8	11·6	5	M	1	118·4
191	2·8	3·2	N	71·1	19·2	13·8		19·3	13·4	69·4	74·2	11·1[a]	11·1	6	M	2	100
192	3·3	3	N	79·8	18·7	14·4	-	19	14·6	76·1	71·4	10	11·1	5	S	1	111
193	3·5	2	M	88·2	18·1	13·8	—	18·3	13·5	73·7	70·1	9	10·9	5	M	1	101
194	2·8	3·8	R	95·4	20·4	15	73·5	20·3	14·8	—	77	9·7	10·8	5 6	M	2	111·3
195	3·4	2·5	M	85·5	18·5	14·2	76·8	18·4	14·3		76·4	9·4	10·4	5	M	3	110·6
196	3·2	2·7	M	82·9	20·2	14·6	.	20·7	14·6	70·5	79	11·4	11·4	6	M	3	100
197	3	3	M	80·6	19·1	14·2	74·3	19·1[a]	14·2[a]	–	79·6	9·7	11·3	5	M	1	116·5
198	3·2	2·8	M	72	19·3	15·2	-	19·5	15·2	77·9	72·5	10·9	12·3	6	S	4	112·8
199	3·3	2·8	R	85·7	20·3	15·8	-	20·5	16·1	78·5	81·7	12·3	12·3	5 6	M	3	100
200	3·8	1·4	M	92·3	19	14·1	74·2	18·7	14	—	82·6	11·4	12	5	S	4	105·3
201	3·7	2·5	M	88·3	20·1	15·8	—	20·3	15·6	76·8	78·1	9·8	12	5	M	1	122·4
202	3·2	2·6	R	80	18·3	14		18·3	14·4	78·7	73·5	9·1	11	5	S	1	120·9
203	3·5	3·3	N	78·7	20·2	14·8	76·5	19·7	14·6	.	80·8	8·2	11·3	5	M	1	137·8
204	2·8	3	R	105·3	18·3	13·6	-	18·9	13·8	73·0	82·2	10	11·3	5	M	1	113
205	3·4	3	M	95·5	20·8	14·6	70·2	20·4	14·5	-	80	9·8	11·5	5	M	1	117·3
206	3·2	2·3	R	88	17·9	14·2	79·3	17·8	14·1	—	74·6	9·3	11·6	5	S	2	124·7
207	3·4	2·6	R	88·5	20·7	15·8	—	20·8	15·8	76	74·8	11·5	12·7	6	M	1	110·4
208	3·3	1·8	R	88·7	18·5	14·7	.	18·9	14·9	78·8	75·9	10·6	12·5	5	M	2	117·9
209	2·5	2·5	M	81·1	18·6	14·9	80·1	18·2	14·1	—	74·4	9·7	10·3	5	S	2	106·2
210	3	3	R	81·4	20·1	15·8[a]	—	20·1	15·8	78·6	77·6	11·6	12	6	S	2	103·4
211	2·8	2·3	R	97·3	19·2[a]	14·5	—	19·2	14·5[a]	75·5	72·4	9·3	11	5	M	1	126·9
212	2·8	2·1	R	82·5	20·8[a]	14·8[a]		20·8	14·8	71·2	83·2	9·1	11·7	5,6	M	2	128·5
213	3·5	2·3	M	79·7	18·3	13·7	74·9	18·3[a]	13·6	-	76·2	8·2	10·7	6/5	M	3	130·5
214	3·1	1·8	R	91	18·7	14·2	76·5	18·6	14·1	-	73·5	9·6	11·6	5	M	1	120·8
215	3	2·6	N	72·6	18·1	13		18·6	12·8	68·8	76·4	9·7	11	5	S	3	113·4
216	2·4	2·5	M	83·8	19·1	14·2	—	19·2	14·2	74·5	73·8	10·8	11·7	6	M	1	108·3
217	3	2·4	N	76·6	19·5	14·5	74·1	19·4	14·7	—	72·2	10·2	10·9	6/5	M	3	106·8
			N-44											4 segs. = 3		1st = 90	
Average †	3·35	2·81	N-M-5	84·26	19·27	14·52	74·6	19·32	14·5	—	75·2	10	11·5	1/5 = 2	S 52	2nd = 59	115·8
Largest ‡	4·5	4·3	M-92	108·9	22·2	16·9	84·1	22·5	16·8	—	90	13·2	13·7	5 = 169	M 146	3rd = 49	174·6
Smallest	2·5	1·1	M-R 4	64	16	11·3	67·8	15·7	11·15	.	63·7	6·3	9·2	5/6 = 11	P 18	4th = 16	86
			R-71											6 = 32	reversed 1	5th = 3	

188. "Male type." 189. Sacral canal open all but 0·1 cm. in middle. 190. Narrow, somewhat asymmetrical pelvis. 191. Six sac. vert., one from coccyx? "Male type." Narrow outlet. 192. Flat pelvis. Small outlet. 194 Transitional vertebra, left sacral, right lumbar type. False promontory. 196. Six sac. vert., one from coccyx (?). 198. Six sac. vert., all sacral character.

* Original number. Specimen not yet permanently catalogued.

199. Transitional vertebra, free on right, one from coccyx. 200. Narrow pelvis. 207. Six sac. vert., all sacral. No false promontory. 210. Six sac. vert., all sacral. No false promontory. 212. Transitional vertebra, right free, but articulates with ilium, left sac. False promontory. 213. Transitional vertebra, left free, right sacralized. False promontory. 216. Six sac. vert., one from coccyx. 217. Transitional vertebra, right sacralized, left free. No false promontory. † False promontory = 18. ‡ Double promontory = 2.

THE INTENSITY OF NATURAL SELECTION IN MAN.

SECOND PAPER.

By E. C. SNOW, M.A., D.Sc.

The present paper is a supplement to the memoir of the same title issued last year*. It is not proposed to give here any account of the work which has been done in the attempt to elicit information on the more difficult subject of the *nature* of selection in man, but only to publish the correlations and regressions obtained by using an alternative measure of environment, and by varying the periods in which the effects of a selective death-rate can be detected.

The adoption of another method of correcting for environment implies not the slightest shaking of my confidence in the adequacy and validity of that employed in the first memoir, but the further work was entered upon because the importance of the subject renders the comparison of the results reached by the use of the various possible methods particularly desirable. The only criticism I have seen of the mode of measuring environment used in the earlier work is that by the Editors of the *Journal of the Royal Statistical Society*. Had there been a tithe of evidence supporting the view adumbrated in that criticism the memoir would have been practically valueless. Fortunately, however, no arguments whatever have been put forward in favour of the view held by the statistical critics, and very cogent facts against that view have already been given†. It is quite beside the point to show that the corrected standard deviation of the total mortality in the two periods considered is only 5 °/ₒ or 6 °/ₒ. That standard deviation is in a number of cases appreciably of the same magnitude as the corresponding measure of dispersion in the earlier of the periods used, and, moreover, is many times its probable error. It matters not whether that standard deviation be 6 °/ₒ or ·06 °/ₒ or 60 °/ₒ of the mean value.

* *Drapers' Company Research Memoirs*, "Studies in National Deterioration," No. VII. Dulau and Co. 1911.

† *Biometrika*, Vol. VIII. p. 456, 1912.

The notions used to indicate environmental conditions for this new work are very simple. Shortly, they consist in making the mortality of a cohort of one sex the measure of the environment for the corresponding cohort of the opposite sex on which it is desired to ascertain the possible effects of selection. Thus if we wish to investigate the selective effect on the male mortality of the third, fourth and fifth years of life of variations in that mortality in the first two years, in addition to fixing the size of the male cohort we fix the size of the corresponding female cohort and also the total female mortality in the first five years of life. We can thus suppose that we are dealing with districts in which the female mortalities up to five years of age for the cohorts born in a particular year are the same. For these districts we find varying male mortalities in each of the periods considered (see Table below), and the mean values of these male mortalities in both periods throughout the whole series of districts can be found. Do the mortalities in the second period of those districts whose mortalities in the first period deviated in the positive direction deviate, on the average, in the positive or negative direction? Districts with the same female environment will possess varying proportions of male weaklings. If these weaklings are killed off in the earlier period, the population which survives to the later one is stronger and likely, therefore, to have a smaller mortality, and this would be indicated by a negative correlation between the mortalities in the two periods (with the proviso dealt with in § XXXIV of the memoir). To the criticism that the total male mortality is highly correlated with the total female mortality, and that by making the latter constant we are practically fixing the former, we can reply by pointing to the considerable standard deviation of the total male mortality when correction is made for constant female mortality (see Table below). Evidence of a more general character, too, can be gathered by turning over the leaves of any of the Registrar-General's valuable Decennial Supplements to his Annual Reports. Pick out a few of the registration districts in which the mortality of one sex for any of the age-groups given is practically the same and compare the mortalities of the other sex among those districts for the same age-group. Quite appreciable variation in the numbers will be found to exist. Not many of such districts can actually be found, but the method of partial correlation essentially consists of a contrivance by which we can for statistical purposes reduce all districts to a constant type.

No method of measuring environment can be theoretically perfect. Districts under the same environmental conditions would not have the same mortality, but the latter would be distributed in some way due to random causes. Of two districts under the same general environment one may in a particular year suffer to a greater extent from epidemics of measles, scarlet fever or summer diarrhœa, and part of the problem of selection consists in ascertaining if these epidemics strike more at the weaker children than at a *random* sample of all children, and this is ascertained by inquiring if the surviving population is more immune in a subsequent period. The two methods which are employed in this paper to measure environment are quite distinct and bear no physical relation to each other. In the one

method the mortality of the cohort corresponding to the cohort whose history is being traced but of the opposite sex, is fixed. In the other, the number of deaths of the same sex and between the same age limits in the period under notice, apart from the deaths occurring within the cohort, is rendered constant. Neither of these, of course, allows for the fact that under a perfectly uniform environment the mortality which is taken to indicate that environment should be distributed in some way due to random causes. But the general similarity which will be shown to exist between the results reached by using the two distinct methods is some justification for the claim that each is a satisfactory approximation to the theoretically best method.

The investigation which is now being described was directed throughout to ascertaining the extent of selection in the mortality of the first two years of life. In the earlier work of which an account is given in the memoir other periods were taken, but later some evidence was adduced to indicate that the first two years of life was a natural interval to adopt, as embracing roughly the whole of the mortality of infancy and overlapping but very little that of childhood. For the second period, on the mortality of which the selective character of that of the earlier period is indicated, the next three years of life are taken both for the English and Prussian data; in the case of the latter, also, the next eight years are employed as a second period. We thus reach results obtained from English and Prussian data by working at the same periods in each case, and the comparison of these results is of interest. The notation employed throughout is:

x_0 = Births of the male or female cohort considered, in, say, year t.

x_1 = Deaths in the cohort in the two years, t and $t+1$.

x_2 = Deaths in the cohort in the next three years, or next eight years.

x_3 = *Remaining* deaths of same sex as cohort in the five years or ten years (see Memoir, § VIII).

x_4 = Deaths in the corresponding cohort of opposite sex in the five years or ten years considered.

x_5 = Births of the corresponding cohort of opposite sex.

Previous experience suggested that the correction for a constant value of x_5 in addition to constant values of x_0 and x_4 would have little effect on the correlation between x_1 and x_2. The first case worked out supported this view; in that (the Prussian male cohort of 1881, dealing with the first ten years of life) the correlation was only altered in the sixth figure by the extra correction, viz. from $-\cdot 944206$ to $-\cdot 944209$. Thus the considerable labour involved in making this further correction is not justified by the extra value it gives to the results, and in all the other cases it was omitted.

The partial standard deviations, correlations and regressions with their probable errors for the various sets of data are given below, the other standard deviations and correlations on which they are based being shown at the end. In the table

Males.

	English Data			Prussian Data (5 Years)		Prussian Data (10 Years)	
	1870 Cohort	1871 Cohort	1872 Cohort	1881 Cohort	1882 Cohort	1881 Cohort	1882 Cohort
$\bar{x}_1$	598	562	538	2270	2424	2270	2424
$\bar{x}_2$	97	101	112	490	514	729	732
$_{03}\sigma_1$	41·1 ± 3·6	44·1 ± 3·8	40·6 ± 3·5	201·1 ± 17·5	161·9 ± 14·1	198·0 ± 17·2	148·4 ± 12·9
$_{04}\sigma_1$	34·3 ± 3·0	39·5 ± 3·4	30·9 ± 2·7	218·9 ± 19·1	225·6 ± 19·7	256·1 ± 22·3	243·7 ± 21·3
$_{03}\sigma_2$	17·3 ± 1·5	13·7 ± 1·2	15·3 ± 1·3	159·6 ± 13·9	176·7 ± 15·4	210·8 ± 18·4	206·4 ± 17·9
$_{04}\sigma_2$	18·5 ± 1·6	14·4 ± 1·3	18·0 ± 1·6	154·5 ± 13·4	157·2 ± 13·7	189·5 ± 16·5	180·4 ± 15·7
$_{03}\sigma_{1+2}$	35·3 ± 3·1	39·8 ± 3·5	40·0 ± 3·5	80·9 ± 7·0	111·7 ± 9·8	78·7 ± 6·9	166·0 ± 14·5
$_{04}\sigma_{1+2}$	31·5 ± 2·7	35·2 ± 3·1	34·9 ± 3·0	96·5 ± 8·4	110·2 ± 9·6	100·6 ± 8·8	105·6 ± 9·2
$_{03}\mathbf{r}_{12}$	− **·521** ± ·090	− **·456** ± ·097	− **·224** ± ·117	− **·922** ± ·018	− **·786** ± ·047	− **·928** ± ·017	− **·605** ± ·078
$_{03}e_{12}$	− ·088	− ·122	− ·109	− ·087	− ·069	− ·096	− ·077
$_{04}\mathbf{r}_{12}$	− **·415** ± ·102	− **·467** ± ·096	− **·058** ± ·123	− **·927** ± ·017	− **·917** ± ·020	− **·944** ± ·013	− **·920** ± ·019
$_{04}e_{12}$	− ·068	− ·104	− ·071	− ·098	− ·108	− ·139	− ·139
$_{03}\mathbf{b}_{21}$	− **·220** ± ·038	− **·141** ± ·030	− **·084** ± ·044	− **·732** ± ·014	− **·857** ± ·051	− **·987** ± ·018	− **·842** ± ·108
$_{04}\mathbf{b}_{21}$	− **·223** ± ·055	− **·171** ± ·035	− **·034** ± ·071	− **·654** ± ·012	− **·639** ± ·014	− **·699** ± ·010	− **·681** ± ·014

Females.

	English Data			Prussian Data (5 Years)		Prussian Data (10 Years)	
	1870 Cohort	1871 Cohort	1872 Cohort	1881 Cohort	1882 Cohort	1881 Cohort	1882 Cohort
$\bar{x}_1$	485	453	437	1914	2004	1914	2004
$\bar{x}_2$	94	98	111	470	505	711	729
$_{03}\sigma_1$	36·4 ± 3·2	33·2 ± 2·9	34·2 ± 3·0	169·7 ± 14·8	134·9 ± 11·8	169·2 ± 14·7	117·4 ± 10·2
$_{04}\sigma_1$	29·8 ± 2·6	24·4 ± 2·1	25·8 ± 2·2	133·0 ± 11·6	124·6 ± 10·9	143·5 ± 12·5	141·2 ± 12·3
$_{03}\sigma_2$	16·0 ± 1·4	18·0 ± 1·6	15·2 ± 1·3	124·6 ± 10·9	155·9 ± 13·6	184·0 ± 16·0	178·7 ± 15·6
$_{04}\sigma_2$	20·2 ± 1·8	22·1 ± 1·9	16·9 ± 1·5	152·2 ± 13·2	158·9 ± 13·8	191·4 ± 16·6	172·8 ± 15·1
$_{03}\sigma_{1+2}$	32·9 ± 2·9	28·6 ± 2·5	35·1 ± 3·1	72·6 ± 6·3	95·0 ± 8·3	98·3 ± 8·6	142·2 ± 12·4
$_{04}\sigma_{1+2}$	33·2 ± 2·9	30·5 ± 2·7	28·3 ± 2·5	115·5 ± 10·0	101·8 ± 8·9	96·9 ± 8·4	106·3 ± 9·2
$_{03}\mathbf{r}_{12}$	− **·429** ± ·100	− **·509** ± ·091	− **·160** ± ·120	− **·921** ± ·018	− **·797** ± ·045	− **·848** ± ·035	− **·608** ± ·077
$_{03}e_{12}$	− ·082	− ·068	− ·092	− ·091	− ·064	− ·093	− ·070
$_{04}\mathbf{r}_{12}$	− **·158** ± ·120	− **·144** ± ·120	− **·172** ± ·119	− **·681** ± ·066	− **·768** ± ·050	− **·876** ± ·029	− **·789** ± ·046
$_{04}e_{12}$	− ·053	− ·041	− ·063	− ·059	− ·058	− ·080	− ·088
$_{03}\mathbf{b}_{21}$	− **·189** ± ·044	− **·277** ± ·049	− **·072** ± ·054	− **·676** ± ·013	− **·921** ± ·052	− **·923** ± ·038	− **·925** ± ·117
$_{04}\mathbf{b}_{21}$	− **·107** ± ·082	− **·131** ± ·109	− **·113** ± ·079	− **·778** ± ·075	− **·979** ± ·063	− **1·169** ± ·039	− **·965** ± ·056

on p. 61, ${}_{03}\sigma_{1+2}$ denotes the partial standard deviation of the total mortality in the sum of the periods considered, ${}_{03}e_{12}$ denotes the expected (partial) correlation if there were no selection (see § XXXIV of the memoir), and ${}_{03}\mathbf{b}_{21}$ denotes the (partial) regression of the mortality of the second period on that of the first.

Dealing first with the results from the English data, we notice that the regressions by the two methods for males are fairly similar, but for females they are, on the whole, smaller by the new method. Having regard to the probable errors we can draw no inferences concerning the differences. The correlations for females, however, are in two of the three cases considerably smaller by the new measure of environment, and this difference appears to be significant. The partial standard deviations by the two methods occasionally show fairly large differences, but in no single case is the disagreement significant. For all the six cases, however, ${}_{04}\sigma_1$ is less than ${}_{03}\sigma_1$ but ${}_{04}\sigma_2$ is greater than ${}_{03}\sigma_2$. The mean of the male regressions by the first method of measuring environment is $-\cdot142$ and by the other $-\cdot143$, the corresponding figures for females being $-\cdot179$ and $-\cdot117$. The mean regressions of the mortality of the 4th and 5th years of life on that of the first three years are $-\cdot085$ and $-\cdot172$ respectively*. Thus, so far as males are concerned, the intensity of selection appears greater when measured by the regression of the mortality of the 3rd, 4th and 5th years on that of the first two years of life than when measured by the regression of the mortality of the 4th and 5th over the first three years, but the same conclusion does not so definitely hold for females. It may be, as was suggested in the memoir, that the age division between infant and child mortality is not the same for females as for males, and the inference is put forward tentatively that the ailments of infancy (as distinct from those of childhood) attack females to a rather greater age than they do males. It will be noticed, too, that the regressions and correlations for the 1872 cohort are smaller than for the other cohorts, and that this is accompanied by the fact that the mortality of that cohort was smaller in the first period. On the whole, so far as the data for the English rural districts are concerned, the adoption of a new measure of environment leads to no alteration of view as regards the existence of selection, nor, roughly, of its numerical intensity.

When we turn to the results from the Prussian data in which the same periods (the first two years and the next three) are used as for the English data, the most marked feature to be noticed is the considerably larger correlations and regressions which are obtained. The mean value of the regressions for males

* I take here the opportunity of correcting a mistake which occurred in connection with the work of the first memoir, through an error in transcribing from the schedules containing the raw data. In § XVIII, in the portion of the Table for the 1872 cohort referring to females, the following alterations should be made:

$$\bar{x}_1 = 487, \quad \sigma_1 = 119\cdot7,$$

$$r_{01} = \cdot812328, \quad r_{12} = \cdot718137, \quad r_{13} = \cdot944379, \quad {}_3r_{01} = +\cdot464859, \quad {}_3r_{12} = -\cdot168670, \quad {}_{103}r_{12} = -\cdot184068.$$

These entail the following corrections in the Table on p. 33 for the same cohort: ${}_{03}\sigma_1 = 34\cdot844$ and partial regression $= -\cdot0591$. These alterations reduce the correlation and regression but make the results more consistent, and necessitate little modification of the conclusions drawn from them.

is $-\cdot794$ by the first method and $-\cdot647$ by the second, while for females the figures are $-\cdot798$ and $-\cdot878$, these comparing with corresponding values for the English districts ranging from $-\cdot12$ to $-\cdot18$. Thus the criterion which we take as the measure of the intensity of selection was for the Prussian cohorts of 1881 and 1882 about five or six times as large as that for the English cohorts of 1870, 1871 and 1872. We can assert with some confidence a considerably greater selective effect of the mortality of the first two years of life on that of the next three in the case of Prussian rural districts than of the English rural districts in the epochs considered, and this fact is concomitant with a far greater stringency of infantile conditions in the former than in the latter. This is seen from the following figures.

		Mean Number of Deaths in First Two Years divided by Mean Number of Births		Mean Number of Deaths in the next Three Years divided by Mean Number of Births minus Deaths in First Two Years	
		Male	Female	Male	Female
England	1870	·185	·157	·037	·036
	1871	·174	·146	·038	·037
	1872	·164	·139	·041	·041
Prussia	1881	·241	·214	·069 (·102)	·067 (·102)
	1882	·260	·228	·075 (·107)	·074 (·107)

The figures in brackets give the corresponding numbers for Prussian districts for the mortality in the eight years following the first two.

The data for English rural districts do not allow us satisfactorily to follow the cohorts beyond their first five years of life, so that we cannot assert that the intensity of selection is *generally* less in England than in Prussia for the populations considered, but we can point out definitely that the effect of selection *in the first five years of life* was much greater in the latter country than in the former. Whether or not the English cohorts make up the leeway at later ages can only be a matter of speculation. We can at present merely state that whereas a district in England which had an excess of 100 male survivors above the mean for all districts at the end of the first two years of life had, on the average, about 14 of these survivors killed off in the next three years, a similar district in Prussia lost more than 70 of the 100 in the same period.

When, for the Prussian data, we come to deal with the results of including the eight years following the first two, we find that the regressions for the 1881 cohort have increased appreciably, but those for the 1882 cohort have not done so. There is nothing incongruous in this, as in one cohort selection might well be felt more in the 3rd, 4th and 5th years than in the other, and in this latter the effect would

then probably come in later years. The tendency of the partial standard deviations for the Prussian data is opposite to that for the English, viz. for males ${}_{04}\sigma_1$ is greater than ${}_{03}\sigma_1$ and ${}_{04}\sigma_2$ is less than ${}_{03}\sigma_2$. For females, however, these are reversed, except for the 1882 cohort in the ten-year period. Another feature of the Prussian results is that the regressions for males are smaller by the new method of measuring environment, but for females the reverse is the case. This is also true for four out of the six examples from English data, and arises chiefly from the differences in the male and female variability in mortality in the second period compared with the first. In the first two years of life the (partial) standard deviation for males is always greater than for females, but for the second period the female (partial) standard deviation is in some cases the larger, the mean mortality in this second period being about the same for the two sexes.

In the memoir (§ XXIV) a short discussion is given to the question of what amount of correlation between the mortalities should be expected if selection were entirely absent. This has been referred to as ${}_{03}e_{12}$ and ${}_{04}e_{12}$ in the present paper. These values are only intended as approximations, and it would undoubtedly be an advantage if by direct correlations we could obviate the use of such corrections. These direct values could be obtained by correlating the mortality *rate* of the first two years of life (based on the number of births) of the cohort with the mortality *rate* in the 3rd, 4th and 5th (or 3rd to 10th) years of life (based on the number of survivors to the age of two), correction being made in some manner for a constant environment rate. This would entail a correlation between such variables as $\frac{x}{y}$ and $\frac{z}{x-y}$ and in my opinion *might* involve an element of 'spurious' correlation, and for this reason alone rates were not used in the memoir. So far as I can understand, however, the critics of the memoir do not hold this opinion, so that to them the corrected correlation between two variates of the above type is probably as satisfactory as a partial correlation of the third order. The employment of such correlations saves considerable labour and requires no discussion of the question of 'expected' correlation if selection were inoperative. Accordingly, for a few cases, the following new variables have been taken:

z_0 = Male or Female Deaths in the first two years of life divided by Male or Female Births,

z_1 = Male or Female Deaths in next three (or eight) years of life divided by number of survivors to the age of two,

z_2 = Total Female or Male (i.e. of opposite sex to z_0 and z_1) Deaths in the whole five (or ten) years divided by Female or Male Births,

and the values of ${}_2r_{01}$ worked out. The statistical constants on which they are based are given below:

	English Rural Districts for 1870		Prussian Rural Districts for 1882			
			Five Years		Ten Years	
	Male	Female	Male	Female	Male	Female
$\bar{z}_0$	·184	·156	·241	·209	·241	·209
$\bar{z}_1$	·037	·036	·069	·069	·100	·103
$\bar{z}_2$	·187	·214	·264	·292	·290	·315
σ_0	·0201	·0184	·0541	·0508	·0541	·0508
σ_1	·0079	·0092	·0263	·0241	·0291	·0253
σ_2	·0232	·0219	·0611	·0653	·0595	·0647
r_{01}	·271543	·463227	·634918	·633825	·596721	·576631
r_{02}	·854400	·877328	·951975	·975109	·937904	·967319
r_{12}	·528884	·567130	·817686	·755164	·811763	·723951
$_2\mathbf{r}_{01}$	**– ·409**	**– ·087**	**– ·814**	**– ·705**	**– ·813**	**– ·707**
$_2\sigma_0$	·0104	·0088	·0166	·0113	·0188	·0129
$_2\sigma_1$	·0067	·0076	·0151	·0158	·0170	·0174
$_2\mathbf{b}_{10}$	**– ·262**	**– ·075**	**– ·744**	**– ·990**	**– ·734**	**– ·958**

If the values of $_2r_{01}$ here are compared with the values of $_{04}r_{12}$ in the earlier tables for the corresponding cases we find very little difference. Thus:

Correlations obtained without using Rates,

$$-·415 \quad -·158 \quad -·917 \quad -·768 \quad -·920 \quad -·789$$

Corresponding Correlations by use of Rates,

$$-·409 \quad -·087 \quad -·814 \quad -·705 \quad -·813 \quad -·707$$

If we reduce the figures in the first of these by the values of $_{04}e_{12}$ we have the sequence

$$-·347 \quad -·105 \quad -·809 \quad -·710 \quad -·781 \quad -·701$$

The agreement between the last two lines of figures is surprising and remarkable. In all cases we get approximately the same numbers as before, a quite unexpected result.

Lest these results should hastily be pointed to as evidence against the possibility of spurious correlation arising when such variables as $\frac{x}{y}$ and $\frac{z}{x-y}$ are correlated, it should be pointed out that the deviations from the mean values of the variables are in some cases considerable, and that the third and fourth powers of those deviations cannot be neglected in comparison with the corresponding powers of the means. Thus the formula which is usually exhibited to show the possibility of spurious correlation does not apply to this case. Whatever the magnitude of the 'spurious' element, if any, involved in the correlations just found, they can only be construed as supporting those previously found and as evidence of the existence of selective mortality in the populations dealt with.

Professor Pearson, also, has pointed out to me another possible mode of attacking the problem. This is to render constant $x_0 - x_1$ instead of x_0, and in

addition to x_3 or x_4, as before. This would fix, not the number of births, but the number of survivors at the end of the first two years. Then the population liable to mortality in the second period would be the same for all districts, but the districts with the larger number of births would be exposed to the possibility of greater mortality in the first period, and therefore to the possibility of greater mortality of the kind which is taken to measure environment (the mortality in the first two years being much greater than in the next three or next eight). Thus an 'expected' (negative) correlation if selection were absent would not arise in the same manner as before, but would probably be entailed in the partial correlation. *A priori*, however, it does not appear that this method would produce a correlation, if selection were inoperative, of the same magnitude as that indicated in § XXIV of the memoir, since the population rendered constant stands in an intermediate position to those at the beginning and end of the periods considered. The correlations under these new conditions have been worked out only for the case of the 1884 Prussian cohort, both male and female. The results differ but very little from those reached before, and must, I think, be taken as supporting the substantial accuracy of the interpretation put upon the earlier ones. If x_5 denote $x_0 - x_1$, the following are the additional correlations:

	Males	Females
σ_5	3117·6	3175·1
r_{51}	·924739	·928746
r_{52}	·847138	·851981
r_{53}	·924711	·942791
r_{54}	·927020	·934079
${}_{53}\mathbf{r}_{12}$	− ·918	− ·747
${}_{54}\mathbf{r}_{12}$	− ·780	− ·806

The values previously found for ${}_{03}r_{12}$ and ${}_{04}r_{12}$ were − ·917 and − ·786 for males, and − ·768 and − ·797 for females. It appears therefore to be of little account whether we make x_0 or $x_0 - x_1$ measure the size of the populations; we should probably, too, get similar results if we put $x_0 - x_1 - x_2$, i.e. the size of the cohorts at the end of our survey, constant.

On the whole, the work of which this paper gives a short account has justified itself by the confirmation and emphasis it gives to the results previously obtained. The general impression received by a study of the results reached by the employment of the new method of measuring environment alone is much the same as that derived from a survey of those by the earlier one, though individual differences of appreciable magnitude occur. Apart from the emphasis it gives to the results of the memoir, the present work has discovered, I think, a significant difference in the operation of selection on the mortality of the first five years of life in Prussian and in English rural districts, and suggests (but, at present, no more than suggests) that there is some differentiation in its effect upon the two sexes. But the existence of a selective death-rate in the general populations dealt with admits of no doubt.

LIST OF STATISTICAL CONSTANTS ON WHICH THE PARTIAL CORRELATIONS AND REGRESSIONS ARE BASED.

English Rural Districts.

	1870 Cohort		1871 Cohort		1872 Cohort	
	Males	Females	Males	Females	Males	Females
$\bar{x}_0$	3227	3090	3226	3114	3291	3150
$\bar{x}_3$	2669	2240	2635	2205	2634	2189
$\bar{x}_4$	579	694	550	663	548	650
σ_0	488·3	467·8	506·6	480·8	520·3	477·0
σ_1	127·7	105·6	130·0	102·6	126·0	107·4
σ_2	27·5	30·6	23·1	28·4	33·0	31·3
σ_3	581·0	508·0	582·3	527·2	605·9	545·1
σ_4	128·6	146·1	120·7	145·8	133·5	153·8
r_{01}	·884086	·856073	·850467	·834253	·793680	·814537
r_{02}	·637792	·643144	·609366	·513438	·716133	·662323
r_{03}	·852236	·831001	·835223	·797029	·807817	·752425
r_{04}	·849348	·881602	·838168	·836549	·824623	·777569
r_{12}	·608255	·688624	·638919	·581274	·804386	·791884
r_{13}	·930579	·925657	·931412	·934695	·945229	·932208
r_{14}	·952680	·958944	·947168	·970562	·969370	·965413
r_{23}	·775362	·845435	·799258	·757490	·886385	·874839
r_{24}	·706226	·749785	·776551	·627105	·838194	·842120
${}_3r_{01}$	+ ·475208	+ ·412648	+ ·362379	+ ·415864	+ ·156490	+ ·475066
${}_4r_{01}$	+ ·467000	+ ·079683	+ ·323442	+ ·169218	− ·040919	+ ·389982
${}_3r_{02}$	− ·069637	− ·199985	− ·176081	− ·229035	+ ·000352	+ ·012762
${}_4r_{02}$	+ ·101579	− ·057213	− ·120807	− ·026164	+ ·080835	+ ·022190
${}_3r_{12}$	− ·490155	− ·464969	− ·482395	− ·546204	− ·221368	− ·135040
${}_4r_{12}$	− ·299955	− ·161886	− ·478055	− ·145896	− ·060726	− ·150136

Prussian Rural Districts (Five-year period).

	1881 Cohort		1882 Cohort	
	Males	Females	Males	Females
$\bar{x}_0$	9407	8917	9297	8793
$\bar{x}_3$	11723	10083	12047	10363
$\bar{x}_4$	2384	2760	2510	2938
σ_0	4586·9	4398·1	4689·5	4504·8
σ_1	1502·2	1276·4	1654·1	1397·1
σ_2	343·2	324·2	375·6	360·7
σ_3	7758·1	6734·5	8044·8	6997·8
σ_4	1558·6	1794·5	1716·2	1981·5
r_{01}	·964196	·963495	·967685	·965286
r_{02}	·844586	·846971	·860460	·869316
r_{03}	·964747	·961005	·967523	·967466
r_{04}	·963378	·970151	·963954	·973849
r_{12}	·817746	·844333	·841894	·864055
r_{13}	·990456	·990191	·995000	·995279
r_{14}	·988301	·994536	·989235	·995822
r_{23}	·884689	·903554	·881938	·901711
r_{24}	·891353	·882250	·906736	·897541
${}_3r_{01}$	+ ·238636	+ ·308402	+ ·198035	+ ·097214
${}_4r_{01}$	+ ·295597	− ·053527	+ ·362344	− ·216609
${}_3r_{02}$	− ·072672	− ·196905	+ ·060222	− ·027966
${}_4r_{02}$	− ·116186	− ·078351	− ·121140	− ·047446
${}_3r_{12}$	− ·910479	− ·919631	− ·756884	− ·795986
${}_4r_{12}$	− ·913825	− ·673466	− ·892589	− ·738506

Prussian Rural Districts (Ten-year period).

	1881 Cohort		1882 Cohort	
	Males	Females	Males	Females
$\bar{x}_0$	9407	8917	9297	8793
$\bar{x}_3$	28277	24650	28137	24480
$\bar{x}_4$	2625	3000	2733	3156
σ_0	4586·9	4398·1	4689·5	4504·8
σ_1	1502·2	1276·4	1654·1	1397·1
σ_2	461·0	445·0	470·7	449·9
σ_3	18056·8	15723·8	18015·8	15645·2
σ_4	1672·8	1901·3	1806·8	2073·1
r_{01}	·966797	·963495	·967684	·965285
r_{02}	·867956	·871859	·895378	·913858
r_{03}	·971578	·968570	·975529	·978221
r_{04}	·969132	·973037	·968885	·977710
r_{12}	·825960	·854617	·860826	·883242
r_{13}	·991230	·991070	·995812	·995626
r_{14}	·983883	·993559	·988210	·994336
r_{23}	·889151	·909771	·890463	·911441
r_{24}	·909835	·902419	·923618	·921165
$_3r_{01}$	+ ·119558	+ ·107783	− ·186994	− ·221147
$_4r_{01}$	+ ·301348	− ·125303	+ ·271614	+ ·161865
$_3r_{02}$	+ ·037627	− ·090243	+ ·266908	+ ·260740
$_4r_{02}$	− ·134827	− ·062670	+ ·005249	− ·308612
$_3r_{12}$	− ·916019	− ·849620	− ·622720	− ·629876
$_4r_{12}$	− ·932723	− ·860001	− ·884404	− ·790688

Also for the 1881 male cohort:

$$r_{05} = \cdot 999225,\quad r_{15} = \cdot 965417,\quad r_{25} = \cdot 866632,$$
$$r_{35} = \cdot 967678,\quad {}_3r_{05} = +\cdot 987800,\quad {}_3r_{15} = +\cdot 295690,$$
$${}_3r_{25} = -\cdot 131819,\quad {}_{03}r_{15} = -\cdot 013340,\quad {}_{03}r_{25} = +\cdot 008833,$$
$${}_{025}r_{12} = -\cdot 944209.$$

ON ERRORS OF RANDOM SAMPLING IN CERTAIN CASES NOT SUITABLE FOR THE APPLICATION OF A "NORMAL" CURVE OF FREQUENCY.

By M. GREENWOOD, Junr.

(*From the Statistical Laboratory of the Lister Institute of Preventive Medicine.*)

I.

Introduction.

THOSE who believe that the more closely a branch of knowledge adapts itself to the principles of quantitative reasoning, the more justly it merits to be ranked as a science, have been gratified by the improved standard adopted in the treatment of medico-statistical results. It is true that even now medico-statistical writers fall short of the attainments regarded as essential in some other departments of natural knowledge, and that a few prominent investigators vaguely denounce "mathematicians"—by which term is to be understood any one trained to employ modern biometric methods—as presumptuous intruders within the sphere of experimental medicine. Despite these obstacles, progress has been marked within recent years and we may have considerable confidence that future discussions as to the value of such procedures as vaccination or the determination of an opsonic index will be conducted with due regard to the claims of exact science.

It is, however, in the nature of things that a reform of this magnitude should be accompanied by certain disadvantages which tend to impede the march of ideas. For instance, reformers may urge that the employment of certain arguments requires for logical validity the application of some specific test. After much discussion, the point is conceded and then the test is in danger of being applied in other and unsuitable instances.

The particular illustration which has prompted these remarks is the employment of some consequences of the current theory of errors of random sampling in certain cases which frequently arise in medical and pathological work. At one time it was customary to base conclusions as to the efficacy of some method of treatment upon short series of cases without any statistical test being employed. A practitioner might find, for example, that of 100 cases of typhoid fever treated without any special precautions as to diet, six had died. Of a subsequent 100, dieted in a particular way, but two succumbed and a conclusion very favourable to the new method of treatment might be ventured. Owing to the partial

permeation of medicine by quantitative methods, it is widely felt that this procedure is questionable and that the "probable error" of the result must be found. The medical writer who has attained this level accordingly refers to a text-book and tests his proportions upon the basis of a "normal" curve of errors with the binomial standard deviation $\sqrt{n\bar{p}\bar{q}}$.

The specialist in mathematical statistics is aware that the time-honoured theory of the "probable error" rests upon certain assumptions of a quite definite character not adequately fulfilled in the imaginary case described. Warnings as to this are given in the better text-books, and are indeed unnecessary for those who care to read the proofs of the usual formulae.

We must, however, bear in mind that not every medical man has either the time or the training requisite for the comprehension of mathematical analysis and many will be inclined to consult a book which, while giving formulae without proofs, contains explicit instructions as to their practical employment. Such a book as, for instance, Professor Davenport's *Statistical Methods*, seems admirably adapted to the needs of the laboratory worker. On p. 14 (2nd edition) he will find the following sentence:—"The probable error of the determination of any value gives the measure of unreliability of the determination; and it should always be found." The statement is commendably clear but, unfortunately, quite incorrect in many cases which come under the notice of the medical inquirer.

The present memoir is an attempt to make the limitations of the process recommended by Professor Davenport arithmetically obvious to the medical reader, and to provide the latter with some assistance in the exceptional cases. To the trained mathematician or biometrician I have nothing to offer which is novel and little which is of interest, while the medical reader may find some difficulties in following every step of the inquiry. I hope these difficulties have been reduced to a minimum, but a risk of falling between two stools has to be faced by any writer dealing with a subject not new in itself but relatively so in its applications. My biometric colleagues will recognise the difficulties of the task, and are alone competent to determine the measure of success or failure achieved.

II.

The chance of an event happening is $\bar{p}$ and of it failing, $\bar{q}$ ($\bar{p}+\bar{q}=1$). What is the "probable error" of $\bar{p}m$ successes in m trials?

In the problem stated the probability for the occurrence of an event and the independence of the happenings are supposed to be known. This *either* means that they have been ascertained by long experience *or* that their values (i.e. the value of $\bar{p}$ and the zero correlation between the results of successive trials) are defined by an hypothesis which we desire to verify. Accordingly the distribution of successes in m trials is given by the expansion of the binomial $(\bar{p}+\bar{q})^m$. If m be moderately large and $\bar{p} \sim \bar{q}$ small, the ordinates of a "normal" curve with Standard Deviation $\sqrt{m\bar{p}\bar{q}}$ are a close approximation to the terms of the binomial

and the "probable error" of $m\bar{p}$ is $\pm \cdot 67449\sqrt{m\bar{p}\bar{q}}$. This is the classical text-book case. Its limitations are obvious. If either $\bar{p}$ or $\bar{q}$ be very small unless m is very large indeed, and for all values of $\bar{p}$ and $\bar{q}$ when m is very small, the normal curve does not approximate closely to the binomial.

Consider this problem. A certain bacillus is stated to occur in the mouths of 2 per cent. of all normal persons. Twenty persons have been examined and the bacillus was isolated from two of them. Is this observation consistent with the truth of the hypothesis?

Let us find the chance that in 20 trials two or more successes would be met with if $\bar{p} = \cdot 02$, $\bar{q} = \cdot 98$.

By direct calculation we find this chance to be about 1 in 17. If we use a "normal" curve with Standard Deviation $\sqrt{20\,(\cdot 02 \times \cdot 98)}$, the chance proves to be rather less than 1 in 25, or the probability determined in this way is only 68 per cent. of the real value. Of course when the number of trials is so small we could not expect a continuous function effectively to represent the binomial expansion, but even for m large the inadequacy of the "normal" curve, in the case of $\bar{p} \sim \bar{q}$ not small, must be insisted upon. I think the best way of making this clear arithmetically is from a consideration of the moment coefficients of the binomial $(\bar{p} + \bar{q})^m$.

With the ordinary notation we have:—

$$\mu_2 = c^2 m\bar{p}\bar{q},$$
$$\mu_3 = c^3 m\bar{p}\bar{q}\,(\bar{p} - \bar{q}),$$
$$\mu_4 = c^4 m\bar{p}\bar{q}\,\{1 + 3\,(m-2)\,\bar{p}\bar{q}\},$$

and $$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \quad \beta_2 = \frac{\mu_4}{\mu_2^2}.$$

In cases like the present, c may be taken as unity.

For a "normal" curve to be a good fit to the binomial, β_1 should be very small and β_2 nearly equal to 3.

Take as an illustration the values of β_1 and β_2 for different values of m where

$$\bar{p} = \cdot 02, \quad \bar{q} = \cdot 98.$$

We obtain:

m	β_1	β_2
100	·4702	3·4502
200	·2351	3·2251
300	·1567	3·1501
400	·1176	3·1126
500	·0940	3·0900
600	·0784	3·0750
700	·0672	3·0643
800	·0588	3·0563
900	·0522	3·0500
1000	·0470	3·0450

From these figures it is plain that even comparatively large values of m do not admit of the binomial being closely approximated to by a "normal" curve.

Since, however, direct evaluation of the terms of the binomial is very tedious when m is at all large, we need a curve which bears the same relation to the skew binomial that the normal curve does to the symmetrical binomial. Such a function was provided years ago by Pearson*, viz. his Skew Curve of Type III,

$$y = y_0 e^{-\gamma x}\left(1 + \frac{\gamma x}{s}\right)^s,$$

where

$$s = \frac{4}{\left(\dfrac{1}{m\bar{p}\bar{q}} - \dfrac{4}{m}\right)} - 1,$$

$$\gamma = \frac{2}{\bar{p} - \bar{q}} \quad \text{(taking the unit of measurement } c = 1 \text{ as before)},$$

$$y_0 = \frac{\gamma s^s \,.\, e^{-s}}{\Gamma(s+1)}.$$

To use this curve with the rapidity possible in the case of a "normal" curve, we need tables not at present published.

In any particular case, however, the curve may be calculated and the area between assigned ordinates approximated to with little labour.

To sum up, we have the following rules for practical work when $\bar{p}$ is known or assumed.

(1) When m is small, say less than 25, the binomial expansion should be directly evaluated.

(2) When m is moderately large and $\bar{p}$ or $\bar{q}$ not small, say not less than ·1, the ordinary method based on the "normal" curve can be trusted.

(3) If m is moderately large and $\bar{p}$ or $\bar{q}$ less than ·1, a skew curve of Type III should be fitted from the momental constants of the binomial and the areas between assigned ordinates estimated with the help of quadrature formulae.

III.

If in n trials an event happened p times and failed q times, what is the probable distribution of successes and failures in m subsequent trials and what are the respective chances of 0, 1, 2, ... *m successes in m trials, it being assumed that the occurrences are independent and that the "universe" of events is indefinitely greater than n + m?*

This problem is of fundamental importance. We note at once that when the last condition is imperfectly fulfilled an important special case may arise, for we then have :—

$\dfrac{n+m}{N}$ finite where N is the number of events comprising the "universe"

* For a recent précis of the relevant facts, see Pearson, K. "On the Curves which are most suitable for describing the Frequency of Random Samples of a Population," *Biometrika*, 1906, Vol. v. p. 172.

or "population" from which the samples come. This problem of drawing from a "limited universe" will not be considered in the present memoir; it has been discussed in the paper of Pearson last cited*.

The class of problem to which attention is now directed may be typified as follows:—

Fifteen "control" rats have been inoculated with a constant dose of a standard culture of plague bacilli and twelve succumbed in a certain time. Ten similar rats have been immunised by a method it is desired to test and five of these died after inoculation with a dose of culture similar to that employed upon the "controls." What is the probability that the deviation from the rate of mortality obtaining among the "controls" is a chance event?

Evidently the methods of pp. 69—72 cannot be used. To state that the *a priori* chance of dying is ·8 is to ignore the fact that the size of the "control" sample does not justify us in assuming that its proportional yield approximates at all closely to that of the whole population.

Let us, then, start from first principles, merely assuming (an assumption based on or supported by the fairly wide practical experience of civilised humanity) that all possible events are, *in the absence of any grounds for inference*, equally likely (Bayes' principle).

On this assumption, we have, by Bayes' Theorem for the chance P_x that the true probability of an event, observed to happen p and fail q times in n trials, is between x and $x + \delta x$:

$$P_x = \frac{x^p (1-x)^q \, dx}{\int_0^1 x^p (1-x)^q \, dx}.$$

A second trial of m being made, the total chance of its yielding r successes and $m-r$ failures is:

$$\frac{m\,!}{r\,!\,(m-r)\,!} \frac{\int_0^1 x^{p+r} (1-x)^{q+m-r} dx}{\int_0^1 x^p (1-x)^q \, dx} \quad \text{.....................(1).}$$

This is, in modern notation, the result contained in the 7th of Condorcet's problems published in his *Essai*, 1785†, but Laplace had, eleven years previously, given the theorem with the omission of the term $\frac{m\,!}{r\,!\,(m-r)\,!}$ (i.e. working on the standard model of an urn from which balls are drawn, he assumed the drawings to have been made in an assigned order).

To Pearson ‡, whose symbols will be used, belongs the credit of emphasizing the enormous statistical value of the theorem. The usual method of treating (1) has

* See Pearson, *op. cit.* pp. 173—5.

† See Todhunter's *History of the Theory of Probability*, p. 383, and for a similar result obtained by a different process in 1795 by Prevost and Lhuilier, *op. cit.* p. 453.

‡ Karl Pearson, "On the Influence of Past Experience on Future Expectation," *Philosophical Magazine*, 1907, p. 365.

been to show that, *under certain conditions*, the probabilities of different values of r can be represented by the ordinates of a "normal" curve*. The nature of the assumptions involved will be placed in the clearest light by the following considerations.

Substituting 0, 1, 2, ... m for r successively in (1), reducing to B and then to Γ functions and finally evaluating each term, we have for the chances of 0, 1, 2, ... m successes in a sample of m after a first sample $n = p + q$:

$$C_0 \left\{1 + \frac{m}{1\,!}\frac{p+1}{q+m} + \frac{m(m-1)(p+1)(p+2)}{2\,!\,(q+m)(q+m-1)} \cdots\right\} \quad \text{............(2)},$$

where
$$C_0 = \frac{\Gamma(q+m+1)\,\Gamma(n+2)}{\Gamma(q+1)\,\Gamma(n+m+2)}.$$

We may notice that, if p and q are both very large as compared with m, (2) reduces to

$$\left(\frac{q}{n}\right)^m \left(1 + \frac{mp}{q1\,!} + \frac{m(m-1)p^2}{2\,!q^2} + \text{etc.}\right)$$
$$= \left(\frac{q}{n}\right)^m \left(1 + \frac{p}{q}\right)^m = (\bar{p} + \bar{q})^m, \text{ where } \bar{p} = \frac{p}{n} \text{ and } \bar{q} = \frac{q}{n} \quad \text{......(2) } \textit{bis}.$$

The conditions for the approximation of this binomial to the "normal" curve have already been noted.

More directly, the approach of (2) to a "normal" form can be examined by treating the series in brackets, which is a hypergeometric series having as parameters

$$\alpha = -m, \quad \beta = p + 1, \quad \gamma = -(q + m), \quad \delta = 1,$$

by the method of moments and then noting the conditions under which the momental constants β_1 and β_2 approximate to the values 0 and 3 respectively. This method was adopted by Pearson who had, several years before the date of the publication last cited, obtained the moment coefficients of a hypergeometric series†.

The results are that:

$$\beta_1 = \frac{(q-p)^2}{m(p+1)(q+1)} \frac{\left(1 + 2\,\dfrac{(m-1)}{(n+4)}\right)^2}{1 + \dfrac{m-1}{n+3}} \quad \text{..................................(3)},$$

$$\beta_2 = 3\left(1 - \frac{2}{m}\right) \frac{1 + \dfrac{m-1}{n+4}\left(1 - \dfrac{8}{m-2} - \dfrac{9}{n+5}\right)}{1 + \dfrac{m-1}{n+3}}$$
$$+ \frac{(n+2)^2}{m(p+1)(q+1)} \frac{1 + 6\,\dfrac{m-1}{n+4}\left(1 + \dfrac{m-2}{n+5}\right)}{1 + \dfrac{m-1}{n+3}} \quad \text{......(4)}.$$

* See, for instance, Czuber's *Wahrscheinlichkeitsrechnung* (1903 Edition), pp. 151 etc.

† Karl Pearson, "On Certain Properties of the Hypergeometrical Series, and on the fitting of such Series to Observation Polygons in the Theory of Chance," *Philosophical Magazine*, 1899, p. 236.

If we write $\bar{p}=\frac{p}{n}$, $\bar{q}=\frac{q}{n}$, $\epsilon=\frac{\bar{p}-\bar{q}}{n+2}$ and if m and n are both absolutely large, we find for (3) and (4)

$$\beta_1=\frac{(\bar{q}-\bar{p})^2}{m(\bar{p}+\epsilon)(\bar{q}-\epsilon)}\frac{\left(1+2\frac{m}{n}\right)^2}{1+\frac{m}{n}} \quad\text{.........................(3 A)},$$

$$\beta_2=3+\frac{1}{m(\bar{p}+\epsilon)(\bar{q}-\epsilon)}\frac{1+6\frac{m}{n}\left(1+\frac{m}{n}\right)}{1+\frac{m}{n}} \quad\text{............(4 A)}.$$

If now m be small *relatively to* n,

$$\beta_1=\frac{(\bar{q}-\bar{p})^2}{m(\bar{p}+\epsilon)(\bar{q}-\epsilon)} \text{ and } \beta_2=3+\frac{1}{m(\bar{p}+\epsilon)(\bar{q}-\epsilon)} \quad\text{.........(5)}.$$

If n be small *relatively to* m,

$$\beta_1=\frac{4(\bar{q}-\bar{p})^2}{n(\bar{p}+\epsilon)(\bar{q}-\epsilon)} \text{ and } \beta_2=3+\frac{6}{n(\bar{p}+\epsilon)(\bar{q}-\epsilon)} \quad\text{............(6)}.$$

After exhibiting these results, Pearson remarks*: "Both forms result—for n or m large and the product of either with $\bar{p}$ and $\bar{q}$ not small—in $\beta_1=0$ and $\beta_2=3$, i.e. in the symmetry and mesokurtosis, which are for practical purposes closely enough represented by the Gaussian curve. But if m and n be commensurable, and either $\bar{p}$ or $\bar{q}$ moderately small, this result by no means follows."

It is accordingly plain that in *all* cases of m and n both small the use of a "normal" curve with S.D. $=\sqrt{m\bar{p}\bar{q}}$ is inappropriate. When $\bar{p}=\bar{q}$ the condition of mesokurtosis is fulfilled and the divergence from "normality" reduces itself to the difference between the Gaussian and Pearson Type II curves. The accompanying table illustrates this in a particular example.

A Second Sample of 10, *after a first Sample of* 100; $p=q=50$.

Comparison of Series with Curves (Totals = 100).

Successes	Hypergeometric Series	Normal Curve S. D. $\sqrt{n\bar{p}\bar{q}}$	Normal Curve S. D. $\sqrt{(n+1)\bar{p}\bar{q}}$	Curve of Type II $y=2\cdot3412\left(1-\frac{x^2}{63\cdot72999}\right)^{10\cdot22177}$
0	·146	·221	·333	·188
1	1·243	1·122	1·408	1·331
2	4·931	4·349	4·843	5·091
3	12·017	11·447	11·702	12·107
4	19·922	20·452	19·728	19·729
5	23·480	24·817	23·972	23·105
6	19·922	20·452	19·728	19·729
7	12·017	11·447	11·702	12·107
8	4·931	4·349	4·843	5·091
9	1·243	1·122	1·408	1·331
10	·146	·221	·333	·188

* *Op. cit.* (1907), pp. 371—2.

A few arithmetical results may now be given.

Let $n = 100$ and $m = 50$.

	From Series (2)	Area* of "Normal" Curve, with S. D. $\sqrt{m\bar{p}\bar{q}}$
$\bar{p} = \cdot 4,\ \bar{q} = \cdot 6$		
Chance of 20—22 Successes	·255†	·271
$\bar{p} = \cdot 1,\ \bar{q} = \cdot 9$		
Chance of 5—7 Successes	·3193	·4739
" 0—2 "	·1311	·1145
$\bar{p} = \cdot 01,\ \bar{q} = \cdot 99$		
Chance of 0—2 Successes	·8938	·9202
" 3—5 "	·1007	·0022

We see how the liability to error increases with $\bar{p} \sim \bar{q}$.

An interesting special case may be discussed here which emphasizes the importance of the problem indicated.

Suppose the first sample has given all successes or all failures, so that $\bar{p}$ or $\bar{q} = 0$, how are we to measure its reliability?

Many unsophisticated users of formulae must have been puzzled by this case, since, construing the formulae *au pied de la lettre*, it would appear that after n successes in n trials, we ought to get m successes in m *with a probable error of* 0!

The paradox vanishes if we consider (2). Put in it $n = p$ and we have

$$\frac{m!\,(n+1)!}{(n+m+1)!}\left\{1 + \frac{n+1}{1!} + \frac{(n+1)(n+2)}{2!} + \text{etc.}\right\} \ldots\ldots\ldots\ldots\ldots (7).$$

From this we see that the ratio of the $(m+1)$th term to the whole sum (i.e. the chance of m successes in m trials) is $\dfrac{n+1}{n+m+1}$; from which we conclude:

(*a*) Only when n is very large as compared with m does the chance of obtaining 100 °/ₒ successes in m trials approach unity.

(*b*) In particular if $n = m$ and both are large, the chance is about ·5.

For instance if we have had 100 °/ₒ successes in 200 trials the chance of getting the same proportion in a subsequent 50 is about 4 to 1. If, on the other hand, $n = 50$ and $m = 200$ it is 1 to 4.

In view of what follows it may be worth noticing that a closed expression for the sum of any number of terms of (7) can readily be given.

* Taking for area corresponding to x successes, the area between the ordinates $x - \cdot 5$ and $x + \cdot 5$.

† Approximate only, obtained by using Stirling's theorem in the expression

$$\frac{m!\,(p+r-1)!\,(q+m-r+1)!}{(m-r+1)!\,p!\,(q+m)!\,(r-1)!},$$

to find the rth term of series (2).

Write in Euler's identity*

$$1-a_1+a_1(1-a_2)+a_1a_2(1-a_3)+\ldots+a_1a_2\ldots a_n(1-a_{n+1})=1-a_1a_2a_3\ldots a_{n+1},$$

$$a_1=\frac{x}{y},\quad a_2=\frac{x+p_1}{y+p_1}\text{ etc.}$$

Multiply by $\dfrac{y}{y-x}$ and subsequently put $y=0$.

We have $$1+\frac{x}{p_1}+\frac{x(x+p_1)}{p_1p_2}+\ldots=\left(\frac{x+p_1}{p_1}\right)\left(\frac{x+p_2}{p_2}\right)\ldots\left(\frac{x+p_n}{p_n}\right)\ \ldots\ldots(8).$$

Putting $$n+1=x,\quad p_1=1,\quad p_2=2,\ \ldots p_m=m,$$

$$1+\frac{n+1}{1}+\frac{(n+1)(n+2)}{2\,!}+\ldots=\prod_{y=1}^{y=m}\left(1+\frac{n+1}{y}\right)\ \ldots\ldots\ldots\ldots(9).$$

Reverting to the general case, we note that for testing the divergence between first and second samples the formula (2) must always be employed when m and n are commensurable and $\bar{p}\sim\bar{q}$ not small. This rule certainly applies to all cases of m and n less than 300 or 400 and $\bar{p}$ (or $\bar{q}$) $<\cdot 1$. If m and n be large the best plan will be to fit to (2) the curve indicated by the momental constants, using its proportional areas (obtained by some convenient quadrature formula) precisely in the manner adopted with the tabled areas of the "normal" curve.

Such a method is, however, not convenient for laboratory workers nor specially appropriate when m is a small number, since in that case the terms of the discontinuous series are not closely represented by a continuous curve.

Evidently what one needs is a tabulation of the series (2) for different values of m, n and p.

Were it possible to obtain a simple formula for the sum of any assigned number of terms of (2), the computation of such a table would be a rapid process.

In the particular case $p=0$ or n, such a formula has been given above. In the general case I have not reached any result† and more widely trained mathematicians, who have kindly allowed me to consult them, do not regard the problem as a simple one.

I therefore fell back upon the method of direct calculation. This is a straightforward but irksome task‡.

* See Chrystal's *Algebra*, Vol. II. p. 392, Ed. 1889.

† Formulae are available in certain types of Hypergeometric Series. *Vide* M. J. M. Hill, "On a Formula for the Sum of a Finite Number of Terms of the Hypergeometric Series when the Fourth Element is equal to Unity," *Proc. Lond. Math. Soc.* 1907, Series 2, Vol. v. p. 335; and 1908, Series 2, Vol. VI. p. 339. The methods of these papers cannot be used in the present case.

‡ Sir Ronald Ross and Mr W. Stott have recently published ("Tables of Statistical Error," *Annals of Tropical Medicine and Parasitology*, Vol. v. No. 3, 1911) a set of tables for the use of laboratory workers. Their tables will be of great service in the cases in which $\bar{p}$ is not less say than 0·1, but are not, I think, available for the class of problem discussed in this paper, since they appear to be based on the "normal" theory of errors. It must be noticed that in an immense number of examples which arise in medical work p will not only be less than 0·1 but less than 0·01 (the prevalence of mental defect in children, albinism, epilepsy, etc. are instances), and for such cases the "normal" treatment is, as pointed out above, inappropriate and often misleading.

For the benefit of those who wish to extend my small tables, it may be worth indicating the arithmetical arrangement which I have found most convenient. I use the following scheme:

$$n, \quad m, \quad p, \quad q, \qquad\qquad\qquad \frac{1}{C_0}$$

$$q + m \qquad\qquad\qquad 100\, C_0$$

Term		Multiplier	
1	(α)	$\dfrac{m\,(p+1)}{(q+m)\,1}$	(a)
$\dfrac{m\,(p+1)}{(q+m)}$	(β)	$\dfrac{(m-1)\,(p+2)}{(q+m-1)\,2}$	(b)
βb	(γ)	$\dfrac{(m-2)\,(p+3)}{(q+m-2)\,3}$	(c)
γc	(δ)	$\vdots$	(d)
$\vdots$	$\vdots$	$\vdots$	$\vdots$

The values of n, m, p, q and $q+m$ are written at the top of the sheet, $\dfrac{1}{C_0}$ and $100\, C_0$ are calculated and written in the right-hand top corner.

Two columns are next formed; the entries in the right-hand column having been made, any given term of the left-hand column is the product of the entries in the columns immediately above it. The entries in the left-hand column are added up and the sum checked by comparing it with $\dfrac{1}{C_0}$. Finally each term is converted into a percentage by multiplying with $100\, C_0$.

In this way, provided one has a mechanical calculator, a series having only a moderately large number of arithmetically significant terms is rapidly evaluated. Still, when all is said, the calculation of a table for values of m and n ranging from say five to a hundred and $\bar{p}$ from 0 to ·1 would need an heroic amount of patience. Even the present admittedly imperfect results have involved the expenditure of some little time and effort* and it was necessary to consider how best to utilise our limited resources.

Having chiefly before my eyes the needs of laboratory workers, I felt sure that the cases of m and n not greater than 25 were of the most importance. Probably in the type of problem alluded to on p. 73 the "control" should be regarded as our n and it is usually possible to arrange the experiments in such a fashion that animals at least equal in number to those specially tested serve as the control. When it is possible to plan a large control, it is usually practicable to fix the

* I desire heartily to thank my assistant, Mr J. W. Brown of the Lister Institute Statistical Department, to whose zealous co-operation in the arithmetical work I am greatly indebted.

number arbitrarily, so that it seemed sufficient to give in tabular form the results of small samples after first samples of 50 and 100 without calculating the intermediate cases.

The next question is as to whether, within the limited field chosen, interpolation can be trusted. Accurate methods of interpolation in the case of double-entry tables are a little complex* and not likely to appeal to the man in the laboratory. What is really material is whether simple interpolation is likely to lead to seriously erroneous conclusions.

I now proceed to some tests.

(1) A first sample of 17 having given 3 successes, required the probability that a second sample of 14 will contain 4 or more successes.

From the tables for $n = 20$, $p = 3$, $m = 10$ and $n = 15$, $p = 3$, $m = 10$, we have for the proportional frequency of 4 or more successes in m trials,

$$\begin{array}{r} 12{\cdot}8062 \\ 23{\cdot}0040 \\ \hline 10{\cdot}1978 \end{array}$$

which gives by simple interpolation for $n = 17$, $p = 3$, $m = 10$,

$$18{\cdot}92488 \qquad (\alpha).$$

Similarly interpolating between the values for $n = 20$, $p = 3$, $m = 15$ and $n = 15$, $p = 3$, $m = 15$, we have for $n = 17$, $p = 3$, $m = 15$,

$$38{\cdot}92372 \qquad (\beta).$$

Interpolating between (α) and (β) we reach for the proportional frequency of 4 or more successes in 14 trials after $n = 17$, $p = 3$, 34·92.

The true value obtained by direct calculation is 35·07601, which gives an error of ·43 °/。 in the interpolated value, a difference of no importance for such purposes as the present. In the accompanying table I have grouped together the results of a number of random trials made in different parts of the table. A perusal of these results leads, I think, to the following conclusions.

(1) For values of m and n ranging in each case up to 25, interpolation, when necessary, gives results of sufficient exactitude for all the purposes likely to be served by such tables.

(2) For greater values of m and n, particularly when the latter is greater than 50, the differences are too great to allow of interpolation and for such values the table can only provide the reader with a general impression (which is often enough sufficient) as to the limits within which possible variations from the proportions

* *Vide* W. Palin Elderton, "Interpolation by Finite Differences (Two Independent Variables)," *Biometrika*, 1903, Vol. II. p. 105; W. Palin Elderton, "Some Notes on Interpolation in n-dimension Space," *ibid.* 1908, Vol. VI. p. 94; also the Introduction to the British Association Tables of $F(r, \nu)$ and $H(r, \nu)$ Functions (issued by *B. A.* 1899, p. 56).

found in the first sample are likely to fall. As remarked above, it was thought that when the first sample exceeded fifty its numerical composition would often be at the choice of the worker. I think, therefore, that these tables are likely to serve most of the objects I had in mind when the work was undertaken, although it is much to be desired that someone will have leisure considerably to extend them. I do not see any immediate prospect of being that person.

The class of problem in which this species of investigation seems desirable has already been described and the reader is perhaps not anxious to see any more arithmetical examples. I may, however, give a single concrete instance of the kind of research in which, I hope, the tables will be of value.

Tests of the Accuracy of Simple Interpolation.

[The True Values are given in brackets.]

Example I. $n=22$, $p=4$, $m=16$.

0—3 Successes	1—3 Successes	3 Successes
57·38 (57·49)	51·27 (51·65)	18·70 (19·14)

Example II. $n=37$, $p=4$, $m=22$.

0—3 Successes	1—2 Successes	3 Successes
62·85 (67·89)	36·47 (39·83)	16·64 (18·88)

Example III. $n=71$, $p=4$, $m=43$.

0—2 Successes	1—3 Successes	3 Successes
46·43 (47·53)	50·60 (56·66)	15·21 (18·25)

Example IV. $n=100$, $p=4$, $m=39$.

0—3 Successes	1—2 Successes	3 Successes
83·84 (84·93)	48·59 (50·78)	13·78 (15·15)

In a paper by Rous*, several experiments of the following kind are detailed.

15 mice† were injected intraperitoneally with a suspension of mouse embryo in normal saline and 11 days later reinjected with the same substance. Ten days after the second injection they were inoculated subcutaneously with a mass made from mouse embryos 1·5 cm. long, and 17 previously untreated mice were inoculated at the same time to serve as a control.

In only one of the 17 control mice was no graft found at the autopsy, but 8 of the treated mice did not "take." If we wish to know whether this difference be an effect of the intraperitoneal inoculations, we may put $n=17$, $p=1$, and ascertain the chance that in $m=15$ there would be 8 or more successes.

From the tables, with interpolation, I find the odds against such a result to be about 260 to 1. In other words it is very likely that the treatment has led to the

* "An Experimental Comparison of Transplanted Tumour and a Transplanted Normal Tissue Capable of Growth," by Peyton Rous, M.D., *Journ. Experimental Medicine*, 1910, Vol. XII. p. 344.

† I take the number stated in the text but can only identify 14 in the corresponding table.

observed result. It may be remarked, however, that had we used a normal curve with S.D. $\sqrt{15 \cdot \frac{1 \times 16}{(17)^2}}$, the odds estimated therefrom would have been enormously greater.

In conclusion, I desire to refer to a subject indirectly related to the topic of this paper and, I think, of importance. We are familiar with such arrangements of material as the following. n_1 persons have been immunised against a certain disease and, having contracted the disease, a_1 have died. Of n_2 not immunised a_2 have died $\left(\frac{a_1}{n_1} < \frac{a_2}{n_2}\right)$. The extent of protection conferred is then estimated by some coefficient of correlation or association. The trustworthiness of the coefficient so calculated is then measured by a comparison between its arithmetical value and that of its standard deviation or "probable error." This process has its limitations. If, for instance, either a_1 or a_2 be zero, Yule's coefficients of association and colligation become unity and their standard deviations are indeterminate.

I would put forward for consideration the possibility that the use of Bayes' theorem might here be of value. Thus:

Let the chance of a_2 or more successes in n_2 after a_1 successes in n_1 be p_2 and the chance of a_1 or less successes in n_1 trials after a_2 successes in n_2 trials be p_1. Then, since either n_1 or n_2 might have been drawn first, a measure of the probability of the observed result will be $\frac{p_1 + p_2}{2}$.

We might indeed adopt a scale of reliability by putting $P = f\left(\frac{p_1 + p_2}{2}\right)$, the function being such that P increases to unity as $\frac{p_1 + p_2}{2}$ diminishes to zero.

I put forward these suggestions with some doubt, but I cannot help feeling sure that in such cases as those I have instanced the ordinary method of testing the reliability of a coefficient of association is a dangerous and possibly misleading artifice*.

In conclusion I desire to express my regret that this paper is so imperfect. The problems treated require mathematical abilities and training not at my command. I have only ventured to write upon the subject because of its practical importance and may, perhaps, venture to entertain the hope that my numerous mistakes of omission and commission will be leniently treated.

* [The probability corresponding to the χ^2 of the fourfold table can be calculated straight away; but the difficulty arises from our not mentally appreciating grades of probability with the readiness we appreciate grades of correlation on a limited scale. EDITOR.]

TABLE. *Percentage Frequency of Successes in a Second Sample "m" after drawing "p" Successes in a First Sample "n".*

	Successes	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$n=6$, $m=5$	0	58·3333	31·8182	15·9091	7·0707		
	1	26·5151	31·8182	26·5151	17·6768		
	2	10·6060	21·2121	26·5151	25·2525		
	3	3·5354	10·6060	18·9394	25·2525		
	4	·8838	3·7879	9·4697	17·6768		
	5	·1263	·7576	2·6515	7·0707		
$n=6$, $m=6$	0	53·8462	26·9231	12·2378	4·8951		
	1	26·9231	29·3706	22·0280	13·0536		
	2	12·2378	22·0280	24·4755	20·3963		
	3	4·8951	13·0536	20·3962	23·3100		
	4	1·6317	6·1189	13·1119	20·3963		
	5	·4079	2·0979	6·1189	13·0536		
	6	·0582	·4079	1·6317	4·8951		
$n=7$, $m=5$	0	61·5385	35·8974	19·5804	9·7902		
	1	25·6410	32·6340	29·3706	21·7560		
	2	9·3240	19·5804	26·1072	27·1950		
	3	2·7972	8·7024	16·3170	23·3100		
	4	·6216	2·7195	6·9930	13·5975		
	5	·0777	·4662	1·6317	4·3512		
$n=7$, $m=6$	0	57·1429	30·7692	15·3846	6·9930		
	1	26·3736	30·7692	25·1748	16·7832		
	2	10·9890	20·9790	25·1748	23·3100		
	3	3·9960	11·1888	18·6480	23·3100		
	4	1·1988	4·6620	10·4895	17·4825		
	5	·2664	1·3986	4·1958	9·3240		
	6	·0333	·2331	·9324	2·7972		
$n=7$, $m=7$	0	53·3333	26·6667	12·3077	5·1282		
	1	26·6667	28·7179	21·5385	13·0536		
	2	12·3077	21·5385	23·4965	19·5804		
	3	5·1282	13·0536	19·5804	21·7560		
	4	1·8648	6·5268	13·0536	19·0365		
	5	·5594	2·6107	6·8531	13·0536		
	6	·1243	·7615	2·6107	6·5268		
	7	·0155	·1243	·5594	1·8648		
$n=8$, $m=5$	0	64·2857	39·5604	23·0769	12·5874	6·2937	
	1	24·7253	32·9670	31·4685	25·1748	17·4825	
	2	8·2418	17·9820	25·1748	27·9720	26·2238	
	3	2·2478	7·1928	13·9860	20·9790	26·2238	
	4	·4495	1·9980	5·2448	10·4895	17·4825	
	5	·0499	·2997	1·0489	2·7972	6·2937	
$n=8$, $m=6$	0	60·0000	34·2857	18·4615	9·2308	4·1958	
	1	25·7143	31·6484	27·6923	20·1398	12·5874	
	2	9·8901	19·7802	25·1748	25·1748	20·9790	
	3	3·2967	9·5904	16·7832	22·3776	24·4755	
	4	·8991	3·5964	8·3916	14·6853	20·9790	
	5	·1798	·9590	2·9370	6·7133	12·5874	
	6	·0200	·1399	·5594	1·6783	4·1958	
$n=8$, $m=7$	0	56·2500	30·0000	15·0000	6·9231	2·8846	
	1	26·2500	30·0000	24·2308	16·1538	9·1783	
	2	11·2500	20·7692	24·2308	22·0280	16·5210	
	3	4·3269	11·5385	18·3566	22·0280	21·4161	
	4	1·4423	5·2448	11·0140	17·1329	21·4161	
	5	·3934	1·8881	5·1399	10·2797	16·5210	
	6	·0787	·4895	1·7132	4·4056	9·1783	
	7	·0087	·0699	·3147	1·0489	2·8846	

TABLE—(*continued*).

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$n=8$, $m=8$	0	52·9412	26·4706	12·3529	5·2941	2·0362	
	1	26·4706	28·2353	21·1765	13·0317	6·7873	
	2	12·3529	21·1765	22·8054	19·0045	12·9576	
	3	5·2941	13·0317	19·0045	20·7322	18·1407	
	4	2·0362	6·7873	12·9576	18·1407	20·1563	
	5	·6787	2·9617	7·2563	12·9000	18·1407	
	6	·1851	1·0366	3·2250	7·2563	12·9576	
	7	·0370	·2633	1·0366	2·9617	6·7873	
	8	·0041	·0370	·1851	·6787	2·0362	
$n=9$, $m=5$	0	66·6667	42·8571	26·3736	15·3846	8·3916	
	1	23·8095	32·9670	32·9670	27·9720	20·9790	
	2	7·3260	16·4835	23·9760	27·9720	27·9720	
	3	1·8315	5·9940	11·9880	18·6480	24·4755	
	4	·3330	1·4985	3·9960	8·1585	13·9860	
	5	·0333	·1998	·6993	1·8648	4·1958	
$n=9$, $m=6$	0	62·5000	37·5000	21·4286	11·5385	5·7692	
	1	25·0000	32·1429	29·6703	23·0769	15·7343	
	2	8·9286	18·5439	24·7253	26·2238	23·6014	
	3	2·7472	8·2418	14·9850	20·9790	24·4755	
	4	·6868	2·8097	6·7433	12·2378	18·3566	
	5	·1249	·6743	2·0979	4·8951	9·4405	
	6	·0125	·0874	·3496	1·0489	2·6224	
$n=9$, $m=7$	0	58·8235	33·0882	17·6471	8·8235	4·0724	
	1	25·7353	30·8824	26·4706	19·0045	11·8778	
	2	10·2941	19·8529	24·4344	23·7557	19·4364	
	3	3·6765	10·1810	16·9683	21·5961	22·6759	
	4	1·1312	4·2421	9·2554	15·1172	20·1563	
	5	·2828	1·3883	3·8873	8·0625	13·6055	
	6	·0514	·3239	1·1518	3·0234	6·4788	
	7	·0051	·0411	·1851	·6170	1·6968	
$n=9$, $m=8$	0	55·5555	29·4118	14·7059	6·8627	2·9412	
	1	26·1438	29·4118	23·5294	15·6863	9·0498	
	2	11·4379	20·5882	23·5294	21·1161	15·8371	
	3	4·5752	11·7647	18·0995	21·1161	20·1563	
	4	1·6340	5·6561	11·3122	16·7969	20·1563	
	5	·5027	2·2624	5·7589	10·7500	16·1250	
	6	·1257	·7199	2·3036	5·3750	10·0782	
	7	·0229	·1645	·6582	1·9197	4·5249	
	8	·0023	·0206	·1028	·3771	1·1312	
$n=9$, $m=9$	0	52·6316	26·3158	12·3839	5·4180	2·1672	
	1	26·3158	27·8638	20·8978	13·0031	6·9659	
	2	12·3839	20·8978	22·2910	18·5759	12·8602	
	3	5·4180	13·0031	18·5759	20·0047	17·5042	
	4	2·1672	6·9659	12·8602	17·5042	19·0955	
	5	·7740	3·2151	7·5018	12·7303	17·1859	
	6	·2381	1·2503	3·6372	7·6382	12·7303	
	7	·0595	·3897	1·4029	3·6372	7·5018	
	8	·0108	·0877	·3897	1·2503	3·2150	
	9	·0011	·0108	·0595	·2381	·7740	
$n=5$, $m=5$	0	54·5454	27·2727	12·1212	4·5454		
	1	27·2727	30·3030	22·7273	12·9870		
	2	12·1212	22·7273	25·9740	21·6450		
	3	4·5454	12·9870	21·6450	25·9740		
	4	1·2987	5·4112	12·9870	22·7273		
	5	·2165	1·2987	4·5454	12·1212		

Percentage Frequencies of Successes in a Second Sample "m".

	Successes	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$
$n=10$, $m=5$	0	68·7500	45·8333	29·4643	18·1318	10·5769	5·7692		
	1	22·9167	32·7381	33·9972	30·2198	24·0385	17·3077		
	2	6·5476	15·1099	22·6648	27·4725	28·8461	26·9230		
	3	1·5110	5·0366	10·3022	16·4835	22·4359	26·9230		
	4	·2518	1·1447	3·0907	6·4103	11·2179	17·3077		
	5	·0229	·1373	·4807	1·2820	2·8846	5·7692		
$n=10$, $m=10$	0	52·3809	26·1905	12·4060	5·5138	2·2704	·8514		
	1	26·1905	27·5689	20·6767	12·9736	7·0949	3·4056		
	2	12·4060	20·6767	21·8930	18·2441	12·7709	7·6625		
	3	5·5138	12·9736	18·2441	19·4604	17·0278	12·5744		
	4	2·2704	7·0949	12·7709	17·0278	18·3377	16·5039		
	5	·8514	3·4056	7·6625	12·5744	16·5039	18·0043		
	6	·2838	1·4190	3·9295	7·8590	12·5030	16·5039		
	7	·0811	·4990	1·6841	4·0826	7·8590	12·5744		
	8	·0187	·1403	·5741	1·6841	3·9295	7·6625		
	9	·0031	·0284	·1403	·4990	1·4190	3·4056		
	10	·0003	·0031	·0187	·0811	·2838	·8514		
$n=15$, $m=5$	0	76·1905	57·1429	42·1053	30·4094	21·4654	14·7575	9·8383	6·3246
	1	19·0476	30·0752	35·0877	35·7757	33·5397	29·5149	24·5958	19·4604
	2	4·0100	10·0251	16·5119	22·3598	26·8318	29·5149	30·2717	29·1906
	3	·6683	2·3588	5·1599	8·9439	13·4159	18·1631	22·7038	26·5369
	4	·0786	·3685	1·0320	2·2360	4·1280	6·8111	10·3199	14·5953
	5	·0049	·0295	·1032	·2752	·6192	1·2384	2·2704	3·8921
$n=15$, $m=10$	0	61·5384	36·9231	21·5385	12·1739	6·6403	3·4783	1·7391	·8238
	1	24·6154	30·7692	28·0936	22·1344	15·8103	10·4348	6·4073	3·6613
	2	9·2308	18·0602	22·9857	23·7154	21·3439	17·2997	12·8146	8·7226
	3	3·2107	8·7565	14·5941	18·9723	20·9694	20·5034	18·0913	14·5376
	4	1·0216	3·6485	7·6619	12·2322	16·3095	18·9958	19·7873	18·6566
	5	·2919	1·3135	3·3874	6·5238	10·3613	14·2469	17·4128	19·1897
	6	·0730	·4032	1·2546	2·8781	5·3965	8·7064	12·4377	15·9914
	7	·0153	·1024	·3795	1·0279	2·2614	4·2644	7·1073	10·6609
	8	·0025	·0203	·0889	·2827	·7269	1·5991	3·1094	5·4516
	9	·0003	·0028	·0145	·0538	·1615	·4146	·9423	1·9383
	10	·0000	·0002	·0012	·0054	·0189	·0565	·1508	·3661
$n=15$, $m=15$	0	51·6129	25·8065	12·4583	5·7842	2·5707	1·0876	·4351	·1631
	1	25·8065	26·6963	20·0222	12·8538	7·4156	3·9155	1·9033	·8512
	2	12·4583	20·0222	20·7638	17·3032	12·4583	7·9941	4·6342	2·4375
	3	5·7842	12·8538	17·3032	17·9953	15·7459	12·0490	8·2152	5·0297
	4	2·5707	7·4156	12·4583	15·7459	16·4305	14·7874	11·7361	8·2991
	5	1·0876	3·9155	7·9941	12·0490	14·7874	15·4916	14·2006	11·5313
	6	·4351	1·9033	4·6342	8·2152	11·7361	14·2006	14·9480	13·8803
	7	·1631	·8512	2·4375	5·0297	8·2991	11·5313	13·8803	14·6968
	8	·0567	·3482	1·1607	2·7663	5·2415	8·3282	11·4309	13·7783
	9	·0181	·1289	·4965	1·3589	2·9443	5·3344	8·3350	11·4309
	10	·0052	·0426	·1882	·5889	1·4548	3·0006	5·3344	8·3282
	11	·0013	·0122	·0618	·2204	·6200	1·4548	2·9443	5·2415
	12	·0003	·0029	·0169	·0689	·2204	·5889	1·3589	2·7664
	13	·0001	·0006	·0037	·0170	·0618	·1882	·4965	1·1607
	14	·0000	·0001	·0006	·0029	·0122	·0426	·1290	·3482
	15	·0000	·0000	·0000	·0003	·0013	·0052	·0181	·0567

TABLE—(*continued*).

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$n=20$	*0*	80·7692	64·6154	51·1538	40·0334	30·9349	23·5695
$m=5$	*1*	16·1538	26·9231	33·3612	36·3940	36·8273	35·3542
	2	2·6923	7·0234	12·1313	17·3305	22·0964	26·0505
	3	·3512	1·2770	2·8884	5·1992	8·1408	11·5780
	4	·0319	·1520	·4333	·9577	1·8090	3·0647
	5	·0015	·0091	·0319	·0851	·1915	·3831
$n=20$	*0*	67·7419	45·1613	29·5884	19·0211	11·9763	7·3700
$m=10$	*1*	22·5806	31·1457	31·7019	28·1795	23·0313	17·6880
	2	7·0078	15·0167	21·1346	24·3861	24·8738	23·2155
	3	2·0022	5·9325	10·8382	15·6071	19·3463	21·5333
	4	·5191	1·9965	4·5521	7·9661	11·7760	15·4158
	5	·1198	·5750	1·5932	3·3250	5·7809	8·8091
	6	·0240	·1398	·4618	1·1335	2·2940	4·0375
	7	·0040	·0278	·1079	·3084	·7210	1·4571
	8	·0005	·0043	·0193	·0636	·1707	·3946
	9	·0000	·0004	·0024	·0089	·0274	·0722
	10	·0000	·0000	·0002	·0006	·0023	·0068
$n=20$	*0*	58·3333	33·3333	18·6275	10·1604	5·3977	2·7859
$m=15$	*1*	25·0000	29·4118	25·4011	19·0508	13·0590	8·3578
	2	10·2941	18·7166	22·2259	21·5090	18·2826	14·1218
	3	4·0553	10·1381	15·5343	18·6411	19·1232	17·4841
	4	1·5207	4·9056	9·3206	13·4987	16·3913	17·4841
	5	·5396	2·1584	4·9495	8·4849	12·0203	14·7942
	6	·1799	·8683	2·3569	4·7138	7·7053	10·8491
	7	·0558	·3190	1·0101	2·3310	4·3590	6·9744
	8	·0159	·1063	·3885	1·0256	2·1795	3·9421
	9	·0041	·0318	·1330	·3989	·9581	1·9511
	10	·0010	·0084	·0399	·1353	·3658	·8362
	11	·0002	·0019	·0102	·0391	·1188	·3041
	12	·0000	·0004	·0022	·0093	·0317	·0907
	13	·0000	·0001	·0004	·0017	·0065	·0209
	14	·0000	·0000	·0000	·0002	·0009	·0033
	15	·0000	·0000	·0000	·0000	·0001	·0003
$n=20$	*0*	51·2195	25·6098	12·4765	5·9099	2·7154	1·2068
$m=20$	*1*	25·6098	26·2664	19·6998	12·7783	7·5427	4·1377
	2	12·4765	19·6998	20·2323	16·8602	12·2839	8·0929
	3	5·9099	12·7783	16·8602	17·3419	15·1742	11·7715
	4	2·7154	7·5427	12·2839	15·1742	15·6340	14·0706
	5	1·2068	4·1377	8·0929	11·7715	14·0706	14·5245
	6	·5172	2·1297	4·9048	8·2768	11·3473	13·3141
	7	·0839	1·0326	2·7589	5·3399	8·3213	11·0186
	8	·0315	·4719	1·4462	3·1817	5·5954	8·3131
	9	·0112	·2030	·7070	1·7554	3·4638	5·7473
	10	·0037	·0819	·3218	·8965	1·9757	3·6474
	11	·0012	·0308	·1358	·4226	1·0362	2·1221
	12	·0003	·0107	·0528	·1829	·4974	1·1274
	13	·0001	·0034	·0188	·0720	·2168	·5429
	14	·0000	·0010	·0060	·0255	·0848	·2345
	15	—	·0003	·0017	·0080	·0293	·0893
	16	—	·0000	·0004	·0022	·0087	·0293
	17	—	—	·0001	·0005	·0022	·0080
	18	—	—	·0000	·0001	·0004	·0017
	19	—	—	—	·0000	·0001	·0003
	20	—	—	—	—	·0000	·0000

TABLE—(*continued*).

	Successes	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$n=25$, $m=5$	0	83·8710	69·8925	57·8421	47·5131	38·7144	31·2693
	1	13·9785	24·1008	30·9868	35·1949	37·2254	37·5232
	2	1·9281	5·1645	9·1813	13·5365	17·8682	21·8885
	3	·2065	·7651	1·7656	3·2488	5·2115	7·6134
	4	·0153	·0736	·2119	·4738	·9064	1·5573
	5	·0006	·0035	·0123	·0329	·0741	·1483
$n=25$, $m=10$	0	72·2222	51·5873	36·4146	25·3798	17·4486	11·8200
	1	20·6349	30·3455	33·1041	31·7248	28·1430	23·6401
	2	5·4622	12·4141	18·6211	23·0261	25·3287	25·6780
	3	1·3242	4·1380	8·0091	12·2806	16·3035	19·5642
	4	·2897	1·1680	2·8032	5·1875	8·1518	11·4125
	5	·0561	·2803	·8119	1·7786	3·2607	5·2673
	6	·0093	·0564	·1933	·4940	1·0451	1·9313
	7	·0013	·0092	·0368	·1086	·2628	·5518
	8	·0001	·0011	·0053	·0179	·0493	·1170
	9	·0000	·0001	·0005	·0020	·0062	·0165
	10	·0000	·0000	·0000	·0001	·0004	·0012
$n=25$, $m=15$	0	63·4146	39·6341	24·3902	14·7625	8·7777	5·1203
	1	23·7805	30·4878	28·8832	23·9392	18·2869	13·1666
	2	8·5366	16·8485	21·8575	23·2742	21·9443	18·9753
	3	2·9204	7·8930	13·1550	17·2894	19·5777	19·9337
	4	·9472	3·2888	6·7654	10·6788	14·2383	16·8191
	5	·2894	1·2403	3·0643	5·6953	8·8100	11·9361
	6	·0827	·4256	1·2381	2·6697	4·7366	7·2943
	7	·0218	·1326	·4477	1·1072	2·2329	3·8807
	8	·0053	·0373	·1444	·4060	·9240	1·8017
	9	·0012	·0094	·0412	·1307	·3337	·7266
	10	·0002	·0020	·0102	·0364	·1038	·2515
	11	·0000	·0004	·0022	·0086	·0272	·0732
	12	·0000	·0001	·0004	·0016	·0058	·0173
	13	·0000	·0000	·0001	·0002	·0009	·0031
	14	—	·0000	·0000	·0000	·0001	·0004
	15	—	—	·0000	·0000	·0000	·0000
$n=25$, $m=20$	0	56·5217	31·4010	17·1278	9·1614	4·7988	2·4579
	1	25·1208	28·5463	23·8993	17·4502	11·7044	7·3738
	2	10·8476	18·9202	21·6231	20·2168	16·6788	12·5733
	3	4·5409	10·8116	15·8218	18·1951	17·9618	15·8820
	4	1·8380	5·6036	10·0864	13·8796	16·0711	16·4186
	5	·7172	2·6897	5·7932	9·3504	12·5094	14·5943
	6	·2690	1·2069	3·0491	5·6861	8·6871	11·4669
	7	·0965	·5082	1·4833	3·1589	5·4604	8·0943
	8	·0330	·2009	·6695	1·6133	3·1317	5·1816
	9	·0107	·0744	·2806	·7592	1·6449	3·0226
	10	·0033	·0257	·1089	·3290	·7916	1·6088
	11	·0009	·0082	·0390	·1308	·3482	·7800
	12	·0002	·0024	·0128	·0475	·1393	·3429
	13	·0001	·0006	·0038	·0156	·0502	·1357
	14	·0000	·0002	·0010	·0046	·0162	·0477
	15	—	·0000	·0002	·0012	·0045	·0147
	16	—	—	·0001	·0002	·0011	·0039
	17	—	—	—	·0000	·0002	·0008
	18	—	—	—	—	·0000	·0001
	19	—	—	—	—	—	·0000
	20	—	—	—	—	—	—

TABLE—(*continued*).

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$n=25$ $m=25$	*0*	50·9804	25·4902	12·4850	5·9824	2·8003	1·2784
	1	25·4902	26·0104	19·5078	12·7285	7·6094	4·2613
	2	12·4850	19·5078	19·9229	16·6024	12·1751	8·1352
	3	5·9824	12·7285	16·6024	16·9713	14·8499	11·6037
	4	2·8003	7·6094	12·1751	14·8499	15·1953	13·6757
	5	1·2784	4·2613	8·1352	11·6037	13·6757	14·0093
	6	·5682	2·2598	5·0451	8·2883	11·1185	12·8418
	7	·2453	1·1411	2·9344	5·4870	8·2990	10·7251
	8	·1027	·5502	1·6103	3·3951	5·7456	8·2555
	9	·0416	·2535	·8365	1·9732	3·7128	5·9003
	10	·0162	·1115	·4118	1·0801	2·2477	3·9335
	11	·0061	·0468	·1921	·5573	1·2771	2·4521
	12	·0022	·0187	·0848	·2709	·6811	1·4304
	13	·0007	·0070	·0353	·1238	·3406	·7802
	14	·0002	·0025	·0138	·0531	·1592	·3971
	15	—	·0008	·0051	·0212	·0693	·1879
	16	—	·0003	·0017	·0079	·0280	·0822
	17	—	·0001	·0005	·0027	·0104	·0330
	18	—	—	·0002	·0008	·0035	·0120
	19	—	—	·0000	·0002	·0010	·0039
	20	—	—	—	·0001	·0003	·0011
	21	—	—	—	—	·0001	·0003
	22	—	—	—	—	—	·0001
	23	—	—	—	—	—	—
	24	—	—	—	—	—	—
	25	—	—	—	—	—	—
$n=50$ $m=5$	*0*	91·0714	82·7922	75·1263	68·0389	61·4967	55·4676
	1	8·2792	15·3319	21·2621	26·1688	30·1454	33·2805
	2	·6133	1·7357	3·2711	5·1311	7·2349	9·5087
	3	·0347	·1335	·3207	·6157	1·0335	1·5848
	4	·0013	·0065	·0192	·0440	·0861	·1517
	5	·0000	·0001	·0005	·0014	·0033	·0066
$n=50$ $m=10$	*0*	83·6065	69·6721	57·8633	47·8869	39·4857	32·4346
	1	13·9344	23·6177	29·9293	33·6048	35·2551	35·3833
	2	2·1256	5·4972	9·4513	13·5019	17·3070	20·6402
	3	·2932	1·0287	2·2503	3·9278	5·9827	8·3080
	4	·0360	·1607	·4296	·8910	1·5803	2·5164
	5	·0039	·0210	·0668	·1614	·3282	·5921
	6	·0003	·0023	·0084	·0233	·0536	·1085
	7	·0000	·0002	·0008	·0026	·0067	·0152
	8	·0000	·0000	·0001	·0002	·0006	·0015
	9	·0000	·0000	·0000	·0000	·0000	·0001
	10	·0000	·0000	·0000	·0000	·0000	·0000
$n=50$ $m=15$	*0*	77·2727	59·4406	45·5092	34·6737	26·2849	19·8214
	1	17·8322	27·8628	32·5066	33·5552	32·3175	29·7321
	2	3·9008	9·2876	14·6804	19·2530	22·6222	24·6927
	3	·8049	2·5965	5·2143	8·3429	11·6306	14·7589
	4	·1558	·6385	1·5643	2·9695	4·8127	6·9910
	5	·0281	·1405	·4083	·9011	1·6718	2·7465
	6	·0047	·0278	·0939	·2371	·4976	·9155
	7	·0007	·0049	·0191	·0544	·1279	·2616
	8	·0001	·0008	·0034	·0109	·0284	·0642
	9	—	·0001	·0005	·0019	·0054	·0132
	10	—	—	·0001	·0003	·0009	·0024
	11	—	—	—	—	·0001	·0003
	12	—	—	—	—	—	·0000
	13, 14, 15	—	—	—	—	—	—

TABLE—(*continued*).

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$n=50$	*0*	71·8310	51·3078	36·4360	25·7195	18·0421	12·5748
$m=20$	*1*	20·5231	29·7437	32·1494	30·7099	27·3365	23·2150
	2	5·6513	12·4661	18·2340	22·1018	23·9720	24·1218
	3	1·4959	4·4655	8·2882	12·2410	15·7316	18·3785
	4	·3796	1·4377	3·2515	5·6902	8·4901	11·3384
	5	·0920	·4247	1·1380	2·3122	3·9438	5·9480
	6	·0212	·1161	·3613	·8391	1·6163	2·7262
	7	·0046	·0295	·1049	·2751	·5926	1·1089
	8	·0010	·0070	·0279	·0820	·1959	·4039
	9	·0002	·0015	·0068	·0222	·0585	·1323
	10	·0000	·0003	·0015	·0055	·0158	·0390
	11	—	·0001	·0003	·0012	·0039	·0103
	12	—	—	·0001	·0002	·0008	·0024
	13	—	—	—	·0000	·0002	·0005
	14	—	—	—	—	·0000	·0001
	15—20	—	—	—	—	—	·0000
$n=50$	*0*	67·1053	44·7368	29·6230	19·4782	12·7149	8·2378
$m=25$	*1*	22·3684	30·2276	30·4346	27·0530	22·3854	17·6525
	2	7·2546	14·9068	20·2898	22·8617	23·0250	21·4900
	3	2·2857	6·3492	10·9546	15·0234	17·9083	19·3831
	4	·6984	2·4592	5·1643	8·3826	11·5877	14·3204
	5	·2066	·8853	2·2004	4·1420	6·5376	9·1130
	6	·0590	·2994	·8629	1·8546	3·3018	5·1406
	7	·0163	·0956	·3146	·7627	1·5167	2·6162
	8	·0043	·0289	·1073	·2904	·6398	1·2147
	9	·0011	·0083	·0343	·1029	·2494	·5181
	10	·0003	·0022	·0103	·0340	·0901	·2038
	11	·0001	·0006	·0029	·0105	·0302	·0741
	12	·0000	·0001	·0008	·0030	·0094	·0249
	13	·0000	·0000	·0002	·0008	·0027	·0077
	14	—	·0000	·0000	·0002	·0007	·0022
	15	—	—	·0000	·0000	·0002	·0006
	16	—	—	—	·0000	·0000	·0001
	17	—	—	—	—	·0000	·0000
	18—25	—	—	—	—	—	·0000
$n=50$	*0*	50·4950	25·2475	12·4963	6·1206	2·9657	1·4210
$m=50$	*1*	25·2475	25·5026	19·1269	12·6198	7·7231	4·4875
	2	12·4963	19·1269	19·3241	16·1034	11·9504	8·1873
	3	6·1206	12·6198	16·1034	16·2729	14·2388	11·2686
	4	2·9657	7·7231	11·9504	14·2388	14·3919	12·9527
	5	1·4210	4·4875	8·1873	11·2686	12·9527	13·0951
	6	·6731	2·5063	5·2821	8·2677	10·6753	12·0038
	7	·3151	1·3552	3·2480	5·7108	8·2014	10·1734
	8	·1457	·7126	1·9185	3·7517	5·9437	8·0780
	9	·0665	·3654	1·0942	2·3606	4·0975	6·0662
	10	·0300	·1831	·6049	1·4298	2·7034	4·3381
	11	·0133	·0898	·3250	·8367	1·7147	2·9694
	12	·0058	·0431	·1699	·4743	1·0490	1·9531
	13	·0025	·0203	·0866	·2610	·6205	1·2381
	14	·0011	·0093	·0431	·1396	·3557	·7582
	15	·0004	·0042	·0209	·0726	·1978	·4493
	16	·0002	·0019	·0099	·0368	·1068	·2580
	17	·0001	·0008	·0046	·0182	·0561	·1437
	18	—	·0003	·0021	·0088	·0286	·0777
	19	—	·0001	·0009	·0041	·0142	·0408
	20	—	—	·0004	·0019	·0069	·0208
	21	—	—	·0002	·0008	·0032	·0103
	22	—	—	·0001	·0004	·0015	·0050
	23	—	—	—	·0002	·0007	·0023
	24	—	—	—	·0001	·0003	·0010
	25	—	—	—	·0000	·0001	·0005
	26	—	—	—	—	·0000	·0002
	27	—	—	—	—	·0000	·0001
	28—50	—	—	—	—	—	·0000

TABLE—(*continued*).

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$
$n=100$, $m=10$	*0*	90·9910	82·7191	75·1302	68·1737	61·8023	55·9719	50·6412	45·7719	41·3280	37·2763
	1	8·2719	15·1778	20·8695	25·4855	29·1520	31·9839	34·0855	35·5510	36·4659	36·9072
	2	·6830	1·8972	3·5107	5·4097	7·4962	9·6874	11·9134	14·1158	16·2472	18·2690
	3	·0506	·1891	·4416	·8243	1·3455	2·0065	2·8031	3·7269	4·7658	5·9052
	4	·0033	·0156	·0442	·0971	·1829	·3098	·4857	·7174	1·0109	1·3708
	5	·0002	·0011	·0036	·0090	·0194	·0368	·0641	·1044	·1609	·2374
	6	·0000	·0001	·0002	·0007	·0016	·0034	·0065	·0115	·0194	·0309
	7	—	·0000	·0000	·0000	·0001	·0002	·0005	·0010	·0017	·0030
	8	—	—	—	—	·0000	·0000	·0000	·0001	·0001	·0002
	9	—	—	—	—	—	—	—	·0000	·0000	·0000
	10	—	—	—	—	—	—	—	—	—	—

		$p=10$	$p=15$	$p=20$	$p=25$	$p=30$	$p=35$	$p=40$	$p=45$	$p=50$
	0	33·5855	19·6056	11·0992	6·0712	3·1945	1·6083	·7697	·3473	·1463
	1	36·9441	33·0200	25·8982	18·5708	12·3788	7·7198	4·5082	2·4580	1·2434
	2	20·1513	26·8727	28·8081	26·8613	22·5639	17·3696	12·3485	8·1229	4·9313
	3	7·1284	13·8698	20·0784	24·1644	25·4567	24·1112	20·8229	16·5036	12·0167
	4	1·8005	5·0127	9·6930	14·9554	19·6711	22·8554	23·9308	22·8255	19·9224
	5	·3376	1·3220	3·3813	6·6468	10·8709	15·4516	19·5797	22·4513	23·4799
	6	·0474	·2571	·8619	2·1464	4·3483	7·5418	11·5470	15·9030	19·9224
	7	·0049	·0363	·1583	·4968	1·2424	2·6232	4·8456	8·0093	12·0167
	8	·0003	·0036	·0200	·0788	·2425	·6221	1·3845	2·7446	4·9313
	9	·0000	·0002	·0015	·0077	·0292	·0908	·2432	·5778	1·2434
	10	—	·0000	·0001	·0004	·0017	·0062	·0199	·0567	·1463

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
$n=100$, $m=5$	*0*	95·2830	90·7457	86·3830	82·1896	78·1607	74·2914	70·5768	67·0123	63·5933	60·3153	57·1739
	1	4·5373	8·7256	12·5800	16·1156	19·3467	22·2874	24·9514	27·3520	29·5020	31·4142	33·1007
	2	·1745	·5083	·9867	1·5956	2·3216	3·1517	4·0737	5·0756	6·1463	7·2749	8·4512
	3	·0051	·0199	·0488	·0957	·1642	·2573	·3780	·5287	·7117	·9287	1·1813
	4	·0001	·0005	·0015	·0034	·0067	·0119	·0197	·0306	·0454	·0649	·0899
	5	·0000	·0000	·0000	·0001	·0001	·0003	·0004	·0008	·0013	·0020	·0030
$n=100$, $m=15$	*0*	87·0690	75·7121	65·7500	57·0221	49·3852	42·7116	36·8873	31·8110	27·3928	23·5527	20·2198
	1	11·3568	19·9243	26·1836	30·5476	33·3684	34·9458	35·5336	35·3456	34·5611	33·3293	31·7739
	2	1·3947	3·7027	6·5459	9·6321	12·7407	15·7096	18·4248	20·8110	22·8233	24·4415	25·6636
	3	·1604	·5730	1·2777	2·2767	3·5456	5·0426	6·7156	8·5076	10·3611	12·2207	14·0361
	4	·0172	·0774	·2091	·4386	·7879	1·2724	1·9006	2·6738	3·5865	4·6273	5·7796
	5	·0017	·0093	·0295	·0715	·1458	·2641	·4381	·6787	·9959	1·3973	1·8884
	6	·0002	·0010	·0037	·0100	·0229	·0461	·0842	·1428	·2278	·3459	·5036
	7	·0000	·0001	·0004	·0012	·0031	·0068	·0137	·0252	·0435	·0711	·1112
	8	—	·0000	·0000	·0001	·0004	·0009	·0019	·0037	·0070	·0122	·0204
	9	—	—	—	·0000	·0000	·0001	·0002	·0005	·0009	·0017	·0031
	10	—	—	—	—	—	·0000	·0000	·0001	·0001	·0002	·0004
	11—15	—	—	—	—	—	—	—	·0000	·0000	·0000	·0000
$n=100$, $m=20$	*0*	83·4711	69·5592	57·8686	48·0604	39·8449	32·9751	27·2403	22·4613	18·4859	15·1848	12·4488
	1	13·9119	23·3813	29·4247	32·8618	34·3491	34·4088	33·4530	31·8036	29·7094	27·3600	24·8976
	2	2·2212	5·6472	9·5567	13·4563	17·0252	20·0718	22·4994	24·2787	25·4270	25·9920	26·0397
	3	·3388	1·1584	2·4716	4·2124	6·2724	8·5261	10·8479	13·1236	15·2562	17·1690	18·8065
	4	·0492	·2122	·5480	1·0993	1·8873	2·9118	4·1535	5·5775	7·1382	8·7832	10·4578
	5	·0068	·0354	·1077	·2490	·4853	·8394	1·3291	1·9649	2·7495	3·6775	4·7356
	6	·0009	·0054	·0191	·0500	·1093	·2099	·3658	·5913	·8994	1·3010	1·8040
	7	·0001	·0008	·0031	·0090	·0219	·0462	·0881	·1547	·2545	·3965	·5898
	8	·0000	·0001	·0004	·0015	·0039	·0090	·0187	·0356	·0630	·1053	·1675
	9	—	·0000	·0001	·0002	·0006	·0016	·0035	·0072	·0137	·0245	·0416
	10	—	—	·0000	·0000	·0001	·0002	·0006	·0013	·0026	·0050	·0091
	11	—	—	—	—	·0000	·0000	·0001	·0002	·0005	·0009	·0017
	12	—	—	—	—	—	—	·0000	·0000	·0001	·0001	·0003
	13—20	—	—	—	—	—	—	—	—	·0000	·0000	·0000

TABLE—(*continued*).

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
$n=100$	*0*	80·1587	64·1270	51·1981	40·7920	32·4330	25·7320	20·3711	16·0915	12·6823	9·9724	7·8232
$m=25$	*1*	16·0317	25·8576	31·2184	33·4361	33·5052	32·1649	29·9575	27·2737	24·3890	21·4922	18·7076
	2	3·1029	7·5681	12·2826	16·5799	20·1031	22·7047	24·3723	25·1757	25·2300	24·6693	23·6306
	3	·5802	1·9024	3·8912	6·3556	9·0661	11·8013	14·3734	16·6391	18·5020	19·9086	20·8423
	4	·1046	·4323	1·0701	2·0562	3·3806	4·9929	6·8150	8·7536	10·7117	12·5970	14·3291
	5	·0182	·0908	·2644	·5855	1·0922	1·8077	2·7378	3·8700	5·1757	6·6134	8·1327
	6	·0030	·0178	·0597	·1501	·3138	·5764	·9606	1·4841	2·1565	2·9790	3·9431
	7	·0005	·0033	·0125	·0351	·0815	·1647	·3000	·5035	·7910	1·1761	1·6693
	8	·0001	·0005	·0024	·0075	·0193	·0426	·0844	·1531	·2589	·4127	·6260
	9	·0000	·0001	·0004	·0015	·0042	·0100	·0215	·0421	·0763	·1299	·2099
	10	—	·0000	·0001	·0003	·0008	·0022	·0050	·0105	·0203	·0369	·0634
	11	—	—	·0000	·0001	·0001	·0004	·0011	·0024	·0049	·0095	·0173
	12	—	—	—	·0000	·0000	·0001	·0002	·0005	·0011	·0022	·0043
	13	—	—	—	—	—	·0000	·0000	·0001	·0002	·0005	·0009
	14	—	—	—	—	—	—	—	·0000	·0000	·0001	·0002
	15—25	—	—	—	—	—	—	—	—	—	·0000	·0000

Successes		$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
$n=100$	*0*	66·8874	44·5916	29·6280	19·6185	12·9456	8·5121	5·5769	3·6405	2·3676	1·5339	·9900
$m=50$	*1*	22·2958	29·9273	30·0284	26·6919	22·1671	17·6113	13·5550	10·1832	7·5029	5·4395	3·8892
	2	7·3322	14·8625	20·0189	22·3956	22·4728	20·9746	18·5789	15·8126	13·0370	10·4710	8·2261
	3	2·3780	6·4708	10·9693	14·8274	17·4789	18·7745	18·8406	17·9434	16·3894	14·4635	12·3988
	4	·7603	2·6038	5·3333	8·4691	11·4896	13·9817	15·7005	16·5656	16·6252	16·0095	14·8876
	5	·2396	·9912	2·3852	4·3589	6·6996	9·1228	11·3492	13·1572	14·4085	15·0512	15·1066
	6	·0743	·3614	1·0008	2·0720	3·5636	5·3759	7·3484	9·2958	11·0430	12·4505	13·4281
	7	·0227	·1271	·3987	·9237	1·7600	2·9173	4·3512	5·9710	7·6559	9·2753	10·7081
	8	·0068	·0433	·1520	·3901	·8167	1·4771	2·3900	3·5398	4·8771	6·3248	7·7895
	9	·0020	·0143	·0557	·1572	·3590	·7044	1·2301	1·9578	2·8874	3·9946	5·2324
	10	·0006	·0046	·0197	·0607	·1504	·3185	·5978	1·0184	1·6022	2·3574	3·2752
	11	·0002	·0014	·0068	·0226	·0603	·1373	·2758	·5012	·8386	1·3088	1·9239
	12	—	·0004	·0022	·0081	·0232	·0566	·1213	·2345	·4161	·6871	1·0664
	13	—	·0001	·0007	·0028	·0086	·0224	·0510	·1046	·1965	·3425	·5601
	14	—	—	·0002	·0009	·0031	·0085	·0206	·0447	·0886	·1627	·2797
	15	—	—	·0001	·0003	·0011	·0031	·0080	·0183	·0382	·0738	·1332
	16	—	—	—	·0001	·0004	·0011	·0030	·0072	·0158	·0320	·0606
	17	—	—	—	—	·0001	·0004	·0011	·0027	·0063	·0133	·0264
	18	—	—	—	—	—	·0001	·0004	·0010	·0024	·0053	·0110
	19	—	—	—	—	—	—	·0001	·0003	·0009	·0020	·0044
	20	—	—	—	—	—	—	—	·0001	·0003	·0008	·0017
	21	—	—	—	—	—	—	—	—	·0001	·0003	·0006
	22	—	—	—	—	—	—	—	—	—	·0001	·0002
	23	—	—	—	—	—	—	—	—	—	—	·0001
	24—50	—	—	—	—	—	—	—	—	—	—	—

ON THE PROBABLE ERROR OF THE CORRELATION COEFFICIENT TO A SECOND APPROXIMATION*.

By H. E. SOPER, M.A.

(1) It is very important in determining whether the coefficient of correlation as found by any particular method differs significantly from the calculated value to know not only its standard deviation but also to have some idea of the nature of the frequency distribution. When the numbers dealt with are large, then, provided r be not nearly ± 1, we may quite legitimately assume a normal distribution and calculate the frequency of r on this basis. But if n be small, or if r have a value near either end of the range, then the usual values for the S.D. of r are not applicable and what is more in the latter case the frequency of r is of a markedly skew character and differs widely from a Gaussian curve. In such case the value of r found from a single sample will most probably be neither the true r of the material nor the mean value of r as deduced from a large number of samples of the same size, but the modal value of r in the given frequency distribution of r for samples of this size. In this paper the following notation will be used:

ρ = correlation coefficient of the material from which the sample is drawn;

$\bar{r}$ = mean value of correlation coefficient for N samples of size n;

$\check{r}$ = modal value of the correlation in the distribution of the values of r as found from N samples of size n;

r = correlation coefficient of any arbitrary sample of size n.

The first question we have to answer is what is likely to be the distribution of the r's. Clearly, when ρ differs from unity, it must be a skew distribution of limited range lying between $+1$ and -1. The general skew curves discussed in *Phil. Trans.*, Vol. 186 A, pp. 343—414, have proved themselves so capable of describing all sorts of types of frequency that one naturally turns to them in the *first* place in the present problem. There appears very little chance of successfully determining—at least for a product-moment table—the distribution of r. We must start with the assumption of a reasonable frequency distribution and justify

* The frequency-distribution of the correlation coefficient in small samples was first discussed by "Student" in his paper in *Biometrika*, Vol. VI. pp. 302–10; he invited further mathematical investigation and to a large extent supplied the impulse and direction to the present paper.

it *a posteriori* by means of experimental samples for given ρ and given n. Now the only type among the skew curves mentioned applicable in the present case is of the form:

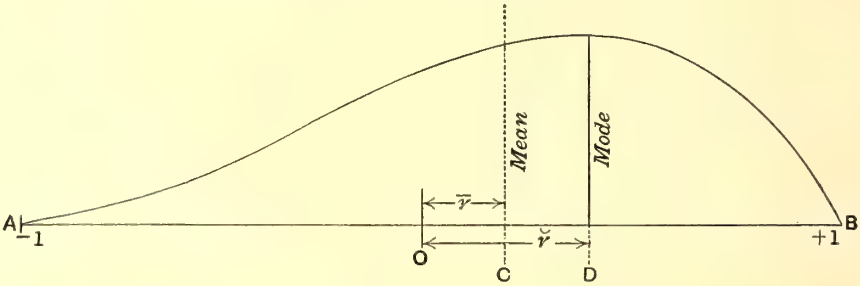
$$y = y_0 \left(1 - \frac{x}{a_1}\right)^{m_1} \left(1 + \frac{x}{a_2}\right)^{m_2} \quad \text{(i)},$$

where, if the origin be at the mode, we must have

$$m_1/a_1 = m_2/a_2 \quad \text{(ii)}$$

Now if we suppose ρ to be positive, we clearly have

$$a_1 = 1 - \check{r}, \qquad a_2 = 1 + \check{r}.$$



Hence from (i) and (ii)

$$y = y_0 \left(1 - \frac{x}{1 - \check{r}}\right)^{m_1} \left(1 + \frac{x}{1 + \check{r}}\right)^{m_2} \quad \text{(iii)},$$

$$\check{r} = (m_2 - m_1)/(m_2 + m_1) \quad \text{(iv)}$$

Now let σ_r denote the standard-deviation of the distribution. Then we easily find (*Phil. Trans., loc. cit.* p. 368)

$$AC = \frac{2(m_2 + 1)}{m_1 + m_2 + 2} = 1 + \bar{r} \quad \text{(v)},$$

$$\sigma_r^2 = \frac{4(m_1 + 1)(m_2 + 1)}{(m_1 + m_2 + 2)^2 (m_1 + m_2 + 3)} \quad \text{(vi)}.$$

Thus $$\bar{r} = (m_2 - m_1)/(m_1 + m_2 + 2) \quad \text{(vii)},$$

$$1 - \bar{r}^2 = 4(m_1 + 1)(m_2 + 1)/(m_1 + m_2 + 2)^2,$$

and $$\sigma_r^2 = (1 - \bar{r}^2)/(m_1 + m_2 + 3) \quad \text{(viii)}.$$

It follows that

$$m_1 + m_2 + 3 = \frac{1 - \bar{r}^2}{\sigma_r^2} = \lambda, \text{ say,}$$

$$m_2 - m_1 = \bar{r}\left\{\frac{1 - \bar{r}^2}{\sigma_r^2} - 1\right\} = \bar{r}(\lambda - 1).$$

Accordingly $$\left.\begin{aligned} m_1 &= \tfrac{1}{2}(\lambda - 1)(1 - \bar{r}) - 1 \\ m_2 &= \tfrac{1}{2}(\lambda - 1)(1 + \bar{r}) - 1 \end{aligned}\right\} \quad \text{(ix)}.$$

Substituting in (iv) we have

$$\left.\begin{aligned} \check{r} &= \bar{r}(\lambda - 1)/(\lambda - 3) \\ \text{and} \quad d &= \check{r} - \bar{r} = 2\bar{r}/(\lambda - 3) \end{aligned}\right\} \quad \text{(x)}.$$

Since $\sigma_r^2 = (1-\bar{r}^2)/\lambda$, and must grow very small as the number in the sample grows large, i.e. λ grows large, we see that $\breve{r}$ and $\bar{r}$ rapidly become equal as the sample increases or the distribution becomes symmetrical.

The value of y_0 can be found from (*Phil. Trans., loc. cit.* p. 369)

$$y_0 = \tfrac{1}{2} N \frac{m_1^{m_1} m_2^{m_2}}{(m_1+m_2)^{m_1+m_2}} \frac{(m_1+m_2+1)\,\Gamma(m_1+m_2+1)}{\Gamma(m_1+1)\,\Gamma(m_2+1)} \quad \text{..........(xi).}$$

The problem of the distribution of r would thus be completely solved, if we knew:

(*a*) r in terms of ρ,

(*b*) σ_r^2 in terms of ρ.

Using Stirling's Theorem we can reduce the expression of y_0 to

$$y_0 = \frac{N}{2\sqrt{2\pi}} \frac{(m_1+m_2+1)\sqrt{(m_1+m_2)}}{\sqrt{m_1 m_2}} e^{\frac{1}{12}\left(\frac{1}{m_1+m_2}-\frac{1}{m_1}-\frac{1}{m_2}\right)}$$

$$= \frac{N}{\sqrt{2\pi}\,\sigma_r} \sqrt{\left(1+\frac{1}{m_1}\right)\left(1+\frac{1}{m_2}\right)} \Big/ \sqrt{\left(1+\frac{3}{m_1+m_2}\right)} \left(1+\frac{1}{m_1+m_2+1}\right) \times e^{\frac{1}{12}\left(\frac{1}{m_1+m_2}-\frac{1}{m_1}-\frac{1}{m_2}\right)}$$

$$= \frac{N}{\sqrt{2\pi}\,\sigma_r}\left\{1+\frac{1}{12}\left(\frac{5}{m_1}+\frac{5}{m_2}-\frac{29}{m_1+m_2}\right)\right\} \quad \text{..........(xii).}$$

This approaches rapidly to the Gaussian value $N/(\sqrt{2\pi}\,\sigma_r)$, if σ_r be at all small and therefore λ and m_1 and m_2 large. For most purposes it is sufficient to take

$$y = \frac{N}{\sqrt{2\pi}\,\sigma_r}\left(1-\frac{x}{1-\breve{r}}\right)^{\frac{1}{2}(\lambda-1)(1-\bar{r})-1}\left(1+\frac{x}{1+\breve{r}}\right)^{\frac{1}{2}(\lambda-1)(1+\bar{r})-1} \quad \text{......(xiii),}$$

where the relation between $\breve{r}$ and $\bar{r}$ is given by (x) and it only remains to determine $\bar{r}$ and σ_r in terms of ρ.

(2) Now the product moment value of the coefficient of correlation, ρ, between two measured characters in any population is defined by

$$\rho = \frac{p_{11} - p_{10}\,p_{01}}{\sqrt{(p_{20}-p_{10}^2)} \times \sqrt{(p_{02}-p_{01}^2)}} \quad \text{..........(xiv),}$$

p_{10}, p_{20} being the first and second moments in respect of the first character, p_{01}, p_{02} those in respect of the second character, and p_{11} being the first product moment, all derived from measures of individuals taken from some arbitrary origins of measurement in the two characters.

If samples of number n are selected at random the moments will have different values p_{10}', p_{20}' etc., and in consequence the coefficient of correlation a different value, r, in any sample, and

$$r = \frac{p_{11}' - p_{10}'\,p_{01}'}{\sqrt{(p_{20}'-p_{10}'^2)} \times \sqrt{(p_{02}'-p_{01}'^2)}} \quad \text{..........(xv).}$$

The mean values of p_{10}', p_{20}' etc. in all samples are p_{10}, p_{20} etc., since the moments are crude and simple averages of individual values. Let dp_{10}, dp_{20}, dp_{01}, dp_{02}, dp_{11} be the deviations of p_{10}', p_{20}', p_{01}', p_{02}', p_{11}' from their means p_{10}, p_{20}, p_{01}, p_{02}, p_{11}. The mean value of r we have called $\bar{r}$. Let dr be the deviation of r from its mean $\bar{r}$, then (xv) becomes

$$\bar{r} + dr = \frac{p_{11} + dp_{11} - (p_{10} + dp_{10})(p_{01} + dp_{01})}{\sqrt{\{p_{20} + dp_{20} - (p_{10} + dp_{10})^2\}} \times \sqrt{\{p_{02} + dp_{02} - (p_{01} + dp_{01})^2\}}} \quad \text{...(xvi)}.$$

Choose the fixed origin of measurement of each character to be the mean of that character in the whole population, then $p_{10} = p_{01} = 0$, and (xvi) becomes

$$\bar{r} + dr = \frac{p_{11} + dp_{11} - dp_{10}dp_{01}}{\sqrt{\{p_{20} + dp_{20} - (dp_{10})^2\}} \times \sqrt{\{p_{02} + dp_{02} - (dp_{01})^2\}}} \quad \text{......(xvii)}.$$

If the distributions and correlations of the deviations of the moments in samples of n are known this is the equation for determining the distribution of the values of the correlation coefficient. The average value of the right-hand side of (xvii) will be $\bar{r}$. The average values of the square, cube etc. of the right-hand expression will give the crude second, third etc. moments of r from which the moments of deviations from mean value of the correlation coefficient can be derived.

Now if (xvii) be expanded in powers and products of the deviations it may be anticipated that the average values of terms of higher order in the deviations are of higher order in $1/n$, and that there is a limit to the number of terms needed to give a required approximation. The approximation sought in the crude moments of r is to terms in $1/n^2$ only, in order that the moments from the mean may be to terms in $1/n^2$, and so that σ_r^2 for instance, which is known* to have the value $(1-\rho^2)^2/n$ to the first approximation for normal frequency, may be further carried to a term in $1/n^2$.

Thus the process for determining $\bar{r}$ is to expand (xvii) and to find and insert in it the average values in samples of n of the various powers and products of the deviations of the moments involved, carrying the process on as far as is necessary to gather in all significant terms as defined above: and a similar process applied to the squares and cubes etc. of (xvii) determines the higher moments of r.

Were the samples sufficiently large these deviations would approximate, as has been shown, to normal distributions, and the known properties of such distributions could be utilised in evaluating the complicated mean, but we are dealing with small samples where the deviations are not so distributed, and it is necessary, in the first place, to evaluate these moments of deviations in terms of the higher moments of the whole distribution without making any assumptions or any approximations within the limits assigned. After this is done the distribution of the two characters will be assumed normal and the results expressed in terms of ρ, the coefficient of correlation of the material examined, and n, the number in the sample, only.

* *Biometrika*, Vol. IX. p. 5 (if $\beta_2 = \beta_2' = 3$).

The method adopted in this paper is that of grade groups. It is well known that if in an indefinitely large population the fraction f fall into a certain grade of a character or combination of grades of two characters, then in taking random samples of n the numbers of this group to be found in such samples follow the binomial distribution of frequency

$$f^n + nf^{n-1}(1-f) + \frac{n(n-1)}{1 \cdot 2} f^{n-2}(1-f)^2 + \ldots + (1-f)^n,$$

and that the mean number is nf and that the deviations from this mean number have moments

$$\text{mean } (df)^2 = \frac{1}{n} f(1-f),$$

$$\text{,, } (df)^3 = \frac{1}{n^2} f(1-f)(1-2f),$$

$$\text{,, } (df)^4 = \frac{3}{n^2} f^2(1-f)^2 + \frac{1}{n^3} f(1-f)(1-6f+6f^2),$$

$$\text{,, } (df)^5 = \frac{10}{n^3} f^2(1-f)^2(1-2f) + \frac{1}{n^4} f(1-f)(1-2f)(1-12f+12f^2),$$

etc., the fourth moment being the last which gives terms in $1/n^2$ [see Pearson, *Phil. Trans.* Vol. 186, p. 347 and *Phil. Mag.* 1899, pp. 240, 241]. Here df is the deviation from mean value f of the frequency of the group in a sample of n.

Moreover if $f_1, f_2, f_3 \ldots$ are the totality of frequencies of the various detached groups into which the population is divided by the graduation (which in our case is a double one) of character the various product moments of the deviations in samples of n may be deduced. These and the above, as far as our approximation needs, are put in one table as follows:

$$\text{mean } (df_1)^2 = \frac{1}{n} f_1(1-f_1) \quad \text{etc.} \quad \ldots\text{(xviii)},$$

$$\text{,, } df_1 \cdot df_2 = -\frac{1}{n} f_1 f_2 \quad \text{,,} \quad \ldots\ldots\text{(xix)},$$

$$\text{,, } (df_1)^3 = \frac{1}{n^2} f_1(1-f_1)(1-2f_1) \quad \text{,,} \quad \ldots\ldots\ldots\text{(xx)},$$

$$\text{,, } (df_1)^2 \cdot df_2 = -\frac{1}{n^2} f_1 f_2 (1-2f_1) \quad \text{,,} \quad \ldots\ldots\text{(xxi)},$$

$$\text{,, } df_1 \cdot df_2 \cdot df_3 = \frac{2}{n^2} f_1 f_2 f_3 \quad \text{,,} \quad \ldots\ldots\text{(xxii)},$$

$$\text{,, } (df_1)^4 = \frac{3}{n^2} f_1^2 (1-f_1)^2 \quad \text{,,} \quad \ldots\ldots\text{(xxiii)},$$

$$\text{,, } (df_1)^3 \cdot df_2 = -\frac{3}{n^2} f_1^2 f_2 (1-f_1) \quad \text{,,} \quad \ldots\ldots\text{(xxiv)},$$

$$\text{,, } (df_1)^2 \cdot (df_2)^2 = \frac{1}{n^2} f_1 f_2 (1 - f_1 - f_2 + 3f_1 f_2) \quad \text{,,} \quad \ldots\ldots\text{(xxv)},$$

$$\text{,, } (df_1)^2 \cdot df_2 \cdot df_3 = -\frac{1}{n^2} f_1 f_2 f_3 (1-3f_1) \quad \text{,,} \quad \ldots\ldots\text{(xxvi)},$$

$$\text{,, } df_1 \cdot df_2 \cdot df_3 \cdot df_4 = \frac{3}{n^2} f_1 f_2 f_3 f_4 \quad \text{,,} \quad \ldots\text{(xxvii)},$$

the last five values being approximate and wanting terms in $1/n^3$ to render them exact.

The method of derivation of the product from the power moments is illustrated in the following example.

$$\text{Mean } (df_1)^2 . df_2 . df_3 = \text{mean } \{(df_1)^2 \times \text{mean } df_2 df_3 \text{ for constant } df_1\}.$$

Now in samples of constant df_1 the number of 1's is $nf_1 + ndf_1$ and of not 1's $n - nf_1 - ndf_1$, amongst which latter restricted population in the whole community the frequency of 2's will be $\frac{f_2}{1-f_1}$ and of 3's $\frac{f_3}{1-f_1}$. Hence the mean number of 2's in such samples will be $(n - nf_1 - ndf_1)\frac{f_2}{1-f_1}$, and of 3's

$$(n - nf_1 - ndf_1)\frac{f_3}{1-f_1},$$

differing from the mean numbers in all samples, nf_2, nf_3, by $-ndf_1\frac{f_2}{1-f_1}$ and $-ndf_1\frac{f_3}{1-f_1}$, and the mean product of the deviations from such means in the restricted samples will be

$$-(n - nf_1 - ndf_1) . \frac{f_2}{1-f_1} . \frac{f_3}{1-f_1}$$

by (xix). It follows that the mean product of the deviations ndf_2, ndf_3, which are measured from the means of all samples, will be

$$-(n - nf_1 - ndf_1) . \frac{f_2}{1-f_1} . \frac{f_3}{1-f_1} + \left(-ndf_1\frac{f_2}{1-f_1}\right)\left(-ndf_1\frac{f_3}{1-f_1}\right),$$

i.e.
$$-n\frac{f_2 f_3}{1-f_1} + ndf_1\frac{f_2 f_3}{(1-f_1)^2} + n^2(df_1)^2\frac{f_2 f_3}{(1-f_1)^2},$$

in samples of constant df_1. Dividing by n^2 we get the value of mean $df_2 . df_3$ for constant df_1 and so obtain finally

$$\text{mean } (df_1)^2 df_2 df_3 = \text{mean} - \frac{(df_1)^2}{n} . \frac{f_2 f_3}{1-f_1} + \frac{(df_1)^3}{n} . \frac{f_2 f_3}{(1-f_1)^2} + (df_1)^4 . \frac{f_2 f_3}{(1-f_1)^2},$$

which by (xviii), (xx) and (xxiii)

$$= -\frac{1}{n^2} f_1 f_2 f_3 + \frac{1}{n^3}\frac{f_1 f_2 f_3 (1-2f_1)}{1-f_1} + \frac{3}{n^2} f_1^2 f_2 f_3$$

$$= -\frac{1}{n^2} f_1 f_2 f_3 (1 - 3f_1),$$

to our approximation.

The other formulae were arrived at in like manner but the process is lengthy and these formulae and the general formulae that follow have been verified by a shorter process, which however being less direct in method is not introduced in this paper.

There is no necessity to take the products of the deviations more than four together, for these do not give terms in $1/n^2$. Did any products, five together for

instance, give terms in $1/n^2$, then mean $(df_1+df_2+df_3+df_4+df_5)^5$ would have such terms, which is contrary to the formula arrived at above for mean $(df)^5$.

Having obtained the various mean products of deviations of group frequencies shown in equations (xviii)—(xxvii) the mean products of deviations of moments, formed by associating with such group frequencies their grade values, follow.

Let a_1, a_2 ... be the values to be assigned to the grades 1, 2 ... in the formation of the moment p (these values will in the present case be the product of one power of one grade of one character with another or the same power of one grade of the second character). Then

$$p = a_1 f_1 + a_2 f_2 + a_3 f_3 + \ldots$$

and if a_1', a_2' ... are the values proper to a second moment, p', in like manner

$$p' = a_1' f_1 + a_2' f_2 + a_3' f_3 + \ldots,$$

and if in random samples of n deviations df_1, df_2 ... in the frequencies lead to deviations dp, dp', in the moments, all deviations being taken from the above universal values which are also the mean values in samples, then

$$dp = a_1 df_1 + a_2 df_2 + a_3 df_3 + \ldots$$
$$dp' = a_1' df_1 + a_2' df_2 + a_3' df_3 + \ldots$$

and so

$$\begin{aligned}\text{mean } dp \,.\, dp' &= \text{mean } [a_1 a_1' (df_1)^2 + a_2 a_2' (df_2)^2 + \ldots + a_1 a_2' df_1 df_2 + a_1' a_2 df_1 df_2 + \ldots] \\ &= \frac{1}{n}[a_1 a_1' f_1 (1-f_1) + a_2 a_2' f_2 (1-f_2) + \ldots - a_1 a_2' f_1 f_2 - a_1' a_2 f_1 f_2 - \ldots] \\ &= \frac{1}{n}[a_1 a_1' f_1 + a_2 a_2' f_2 + \ldots - (a_1 f_1 + a_2 f_2 + \ldots)(a_1' f_1 + a_2' f_2 + \ldots)].\end{aligned}$$

If then p is the u, v moment defined by

$$p_{uv} = a_1^u b_1^v f_{11} + a_1^u b_2^v f_{12} + a_2^u b_2^v f_{22} + \ldots$$

obtained by summing the products of the group frequencies f by the uth power of the grade value a of the first character and the vth power of the grade value b of the second character in that group; and p' is the u', v' moment defined by

$$p_{u'v'} = a_1^{u'} b_1^{v'} f_{11} + a_1^{u'} b_2^{v'} f_{12} + a_2^{u'} b_2^{v'} f_{22} + \ldots,$$

it follows that the first term in the above square brackets is

$$a_1^{u+u'} b_1^{v+v'} f_{11} + a_1^{u+u'} b_2^{v+v'} f_{12} + \ldots$$

or $p_{u+u'\,v+v'}$, and the general formula for the mean products two together of deviations of grade moments is*

$$\text{mean } dp_{uv} \,.\, dp_{u'v'} = \frac{1}{n}[p_{u+u'\,v+v'} - p_{uv}\,p_{u'v'}] \quad \ldots\ldots\ldots\ldots\text{(xxviii)}.$$

* See W. F. Sheppard, "On the application of the theory of error to cases of normal distribution and correlation," *Phil. Trans.* 1899 (192 A), in which paper (p. 127) are given formulae for the mean products, two together, of errors of moments calculated from the means of samples. In the present paper, it should be noted, the moments of the samples are crude, being calculated, not from the means of the samples, but from the mean values of the measured characters in the whole population; and dp is the deviation in the value of the crude moment in any particular sample from its mean value in all samples, which is mean $a_1(f_1+df_1)+a_2(f_2+df_2)+\ldots=a_1f_1+a_2f_2+\ldots=p$ or the moment in the whole population. This latter is a true moment, the general means having been taken as the origin of measurement.

It will be observed that nothing in the proof prevents u', v' from having the same values as u, v and the formula is true for any second order moment whether power or product.

In like manner if p, p', p'' are any three moments of the material sampled we have the equations of deviations

$$dp = a_1 df_1 + a_2 df_2 + a_3 df_3 + \dots$$
$$dp' = a_1' df_1 + a_2' df_2 + a_3' df_3 + \dots$$
$$dp'' = a_1'' df_1 + a_2'' df_2 + a_3'' df_3 + \dots$$

giving

$$\begin{aligned}\text{mean } dp \,.\, dp' \,.\, dp'' = \text{mean } [&a_1 a_1' a_2'' (df_1)^3 + \dots \text{ all grades} \\ &+ (a_1 a_1' a_2'' + a_1' a_1'' a_2 + a_1 a_1'' a_2')(df_1)^2 df_2 \\ &\quad + (a_1 a_2' a_2'' + a_1' a_2 a_2'' + a_1'' a_2 a_2') df_1 \,.\, (df_2)^2 + \dots \text{ all pairs} \\ &+ (a_1 a_2' a_3'' + a_1 a_2'' a_3' + a_1' a_2 a_3'' \\ &\quad + a_1' a_2'' a_3 + a_1'' a_2 a_3' + a_1'' a_2' a_3)\, df_1 \,.\, df_2 \,.\, df_3 + \dots \text{all triads}],\end{aligned}$$

and inserting the values from equations (xx), (xxi), (xxii),

$$\begin{aligned}\text{mean } dp \,.\, dp' \,.\, dp'' = \frac{1}{n^2}[&a_1 a_1' a_1'' f_1 (1 - f_1)(1 - 2f_1) + \dots \text{ all grades} \\ &- (a_1 a_1' a_2'' + a_1' a_1'' a_2 + a_1 a_1'' a_2') f_1 f_2 (1 - 2f_1) \\ &- (a_1 a_2' a_2'' + a_1' a_2 a_2'' + a_1'' a_2 a_2') f_1 f_2 (1 - 2f_2) - \dots \text{ all pairs} \\ &+ (a_1 a_2' a_3'' + a_1 a_2'' a_3' + a_1' a_2 a_3'' \\ &\quad + a_1' a_2'' a_3 + a_1'' a_2 a_3' + a_1'' a_2' a_3)\, 2 f_1 f_2 f_3 + \dots \text{ all triads}],\end{aligned}$$

and collecting terms of first, second and third degree in f's and suitably commuting the f's and a's this is seen to be

$$\begin{aligned}&= \frac{1}{n^2}[(a_1 a_1' a_1'' f_1 + \dots) \\ &- (a_1 a_1' f_1 + \dots)(a_1'' f_1 + \dots) - (a_1 a_1'' f_1 + \dots)(a_1' f_1 + \dots) - (a_1' a_1'' f_1 + \dots)(a_1 f_1 + \dots) \\ &\quad + 2(a_1 f_1 + \dots)(a_1' f_1 + \dots)(a_1'' f_1 + \dots)],\end{aligned}$$

the sums being for all grades.

If then p, p' have the double grade values $p_{uv}, p_{u'v'}$ previously assigned and p'' stands in the same way for $p_{u''v''}$ where

$$p_{u''v''} = a_1^{u''} b_1^{v''} f_{11} + a_1^{u''} b_2^{v''} f_{12} + a_2^{u''} b_2^{v''} f_{22} + \dots$$

there results the general formula for the mean products three together of deviations in moments as follows

$$\begin{aligned}\text{mean } dp_{uv} \,.\, dp_{u'v'} \,.\, dp_{u''v''} = \frac{1}{n^2}[&p_{u+u'+u''\,v+v'+v''} - p_{u+u'\,v+v'}\, p_{u''v''} \\ &- p_{u+u''\,v+v''}\, p_{u'v'} - p_{u'+u''\,v'+v''}\, p_{uv} + 2 p_{uv} p_{u'v'} p_{u''v''}] \quad \dots\dots\text{(xxix)},\end{aligned}$$

where, as before, the values of the suffixes may be any the same and the formula gives power moments equally well with the product moments of the deviations.

Precisely the same process evaluates the mean products four together. We shall have, putting in representative terms only of each series, and using equations (xxiii), (xxiv), (xxv), (xxvi), (xxvii),

$$\begin{aligned}\text{mean } dp\,.\,dp'\,.\,dp''\,.\,dp''' = \frac{1}{n^2}[&a_1a_1'a_1''a_1'''3f_1^2(1-f_1)^2 + \dots\\ &- a_1a_1'a_1''a_2'''3f_1^2(1-f_1)f_2 + \dots\\ &+ a_1a_1'a_2''a_2'''f_1f_2(1-f_1-f_2+3f_1f_2) + \dots\\ &- a_1a_1'a_2''a_3'''f_1f_2f_3(1-3f_1) - \dots\\ &+ a_1a_2'a_3''a_4'''f_1f_2f_3f_4 + \dots]\end{aligned}$$

$$\begin{aligned}= \frac{1}{n^2}[&(a_1a_1'f_1+\dots)(a_1''a_1'''f_1+\dots)\\ &+(a_1a_1''f_1+\dots)(a_1'a_1'''f_1+\dots)\\ &+(a_1a_1'''f_1+\dots)(a_1'a_1''f_1+\dots)\\ &-(a_1a_1'f_1+\dots)(a_1''f_1+\dots)(a_1'''f_1+\dots)\\ &-(a_1a_1''f_1+\dots)(a_1'f_1+\dots)(a_1'''f_1+\dots)\\ &-(a_1a_1'''f_1+\dots)(a_1'f_1+\dots)(a_1''f_1+\dots)\\ &-(a_1'a_1''f_1+\dots)(a_1f_1+\dots)(a_1'''f_1+\dots)\\ &-(a_1'a_1'''f_1+\dots)(a_1f_1+\dots)(a_1''f_1+\dots)\\ &-(a_1''a_1'''f_1+\dots)(a_1f_1+\dots)(a_1'f_1+\dots)\\ &+3(a_1f_1+\dots)(a_1'f_1+\dots)(a_1''f_1+\dots)(a_1'''f_1+\dots)]\end{aligned}$$

on collecting terms and rearranging the associations of f's and a's as before. And putting into factors this

$$\begin{aligned}=\frac{1}{n^2}[&\{(a_1a_1'f_1+\dots)-(a_1f_1+\dots)(a_1'f_1+\dots)\}\{(a_1''a_1'''f_1+\dots)-(a_1''f_1+\dots)(a_1'''f_1+\dots)\}\\ &+\{(a_1a_1''f_1+\dots)-(a_1f_1+\dots)(a_1''f_1+\dots)\}\{(a_1'a_1'''f_1+\dots)-(a_1'f_1+\dots)(a_1'''f_1+\dots)\}\\ &+\{(a_1a_1'''f_1+\dots)-(a_1f_1+\dots)(a_1'''f_1+\dots)\}\{(a_1'a_1''f_1+\dots)-(a_1'f_1+\dots)(a_1''f_1+\dots)\}].\end{aligned}$$

And so again if the material is double graded and p_{uv}, $p_{u'v'}$, $p_{u''v''}$, $p_{u'''v'''}$ are any four moments involving products of powers of both grades, the general formula for the mean products four together of the deviations of such moments in samples of n becomes

$$\begin{aligned}&\text{mean } dp_{uv}\,.\,dp_{u'v'}\,.\,dp_{u''v''}\,.\,dp_{u'''v'''}\\ &\quad=\frac{1}{n^2}[(p_{u+u'\,v+v'}-p_{uv}\,p_{u'v'})(p_{u''+u'''\,v''+v'''}-p_{u''v''}\,p_{u'''v'''})\\ &\qquad+(p_{u+u''\,v+v''}-p_{uv}p_{u''v''})(p_{u'+u'''\,v'+v'''}-p_{u'v'}p_{u'''v'''})\\ &\qquad+(p_{u+u'''\,v+v'''}-p_{uv}p_{u'''v'''})(p_{u'+u''\,v'+v''}-p_{u'v'}\,p_{u''v''})]\quad\dots\dots\dots(\text{xxx}),\end{aligned}$$

and it is to be recalled that this formula omits terms in $\frac{1}{n^3}$, not wanted to the degree of approximation laid down.

Comparing (xxx) and (xxviii) it appears that the mean values four together are equal, within our degree of approximation, to the sum of the products of the mean values two together of the complementary pairs making the four, the division being possible in three ways.

$$\begin{aligned}\text{mean } dp_{uv}dp_{u'v'}dp_{u''v''}dp_{u'''v'''} &= \text{mean } dp_{uv}dp_{u'v'} \times \text{mean } dp_{u''v''}dp_{u'''v'''}\\ &+ \text{mean } dp_{uv}dp_{u''v''} \times \text{mean } dp_{u'v'}dp_{u'''v'''}\\ &+ \text{mean } dp_{uv}dp_{u'''v'''} \times \text{mean } dp_{u'v'}dp_{u''v''} \quad \ldots\text{(xxxi)}.\end{aligned}$$

It is unnecessary to find the general formula for the mean products of deviations five together, which by p. 96 will contribute nothing within our approximation, and formulae (xxviii), (xxix) and (xxx) applied to the expansions of (xvii) and its powers are sufficient to evaluate the general formulae for the mean and moments of deviations of r as far as terms in $1/n^2$.

It is not proposed to exhibit these general formulae for moments of deviations of r in terms of the higher moments of the given distribution at length, but to proceed at once to the simpler case of a normal distribution in the two correlated characters and reduce the higher moments to second moments and the coefficient of correlation, ρ, as such distributions, it is well known, admit. In order to reduce in this way the values of the various mean products at the same time that they are evaluated by the formulae (xxviii), (xxix), (xxx), the necessary formulae of reduction are next obtained.

The expression for dr involves dp_{10}, dp_{20}, dp_{01}, dp_{02} and dp_{11}, and (xxix) shows that we shall require to reduce p_{60}, $p_{51} \ldots p_{50} \ldots p_{40} \ldots p_{30} \ldots$ in the above way.

Now it is well known that in normal distributions, following the Gaussian law of frequency, the odd moments from the mean in either character vanish and the even moments are derived from the second moment by a simple formula of reduction, from which there results that

$$p_{50} = p_{30} = p_{10} = 0, \qquad p_{05} = p_{03} = p_{01} = 0,$$

$$p_{40} = 3p_{20}^2, \quad p_{60} = 5 \cdot 3p_{20}^3, \quad p_{04} = 3p_{02}^2, \quad p_{06} = 5 \cdot 3p_{02}^3.$$

And, utilising these results, the higher product moments of normal distributions in two characters may be derived from the first product moment and the second moments by two well-known properties of the Gaussian surface.

If x, y are deviations from their mean value of two normally correlated characters the mean value of y for a given x is $\frac{p_{11}}{p_{20}}x$, and if y' is the deviation

of y from its mean value in the array the distribution of y' is normal; its second moment is $p_{02}-\frac{p_{11}^2}{p_{20}}$; and its higher moments follow the same laws of reduction as above. Hence

$$\begin{aligned}
p_{uv} &= \text{mean } x^u y^v \\
&= \text{mean } \{x^u \times \text{mean } y^v \text{ for given } x\} \\
&= \text{mean } \left\{x^u \times \text{mean } \left(\frac{p_{11}}{p_{20}}x + y'\right)^v \text{ for given } x\right\} \\
&= \text{mean } \left\{x^u \times \text{mean } \left(\frac{p_{11}^{\,v}}{p_{20}^{\,v}}x^v + v\,.\,\frac{p_{11}^{\,v-1}}{p_{20}^{\,v-1}}x^{v-1}y' \right.\right. \\
&\qquad \left.\left. + \frac{v\,.\,(v-1)}{1\,.\,2}\,.\,\frac{p_{11}^{\,v-2}}{p_{20}^{\,v-2}}x^{v-2}y'^2 + \dots\right) \text{ for given } x\right\},
\end{aligned}$$

and so remembering that mean y', mean y'^3, etc. vanish, and mean y'^4, mean y'^6, etc. reduce by the above formulae we have

$$\begin{aligned}
p_{uv} &= \left(\frac{p_{11}}{p_{20}}\right)^v p_{u+v\,0} + \frac{v\,(v-1)}{1\,.\,2}\left(\frac{p_{11}}{p_{20}}\right)^{v-2} p_{u+v-2\,0} \\
&\quad + \frac{v\,(v-1)\,(v-2)\,(v-3)}{1\,.\,2\,.\,3\,.\,4}\left(\frac{p_{11}}{p_{20}}\right)^{v-4} p_{u+v-4\,0}\left(p_{02}-\frac{p_{11}^2}{p_{20}}\right) \\
&\quad + \frac{v!}{6!\,(v-6)!}\left(\frac{p_{11}}{p_{20}}\right)^{v-6} p_{u+v-6\,0}\,.\,3\left(p_{02}-\frac{p_{11}^2}{p_{20}}\right)^2 + \dots \qquad \text{.........(xxxii).}
\end{aligned}$$

It follows that if $u+v$ is odd p_{uv} is zero, that is

$$p_{30}=p_{21}=p_{12}=p_{03}=0,$$
$$p_{50}=p_{41}=p_{32}=p_{23}=p_{14}=p_{05}=0.$$

If $u+v$ is even it is convenient to divide by suitable powers of p_{20} and p_{02}, and putting ρ for $\frac{p_{11}}{\sqrt{p_{20}}\sqrt{p_{02}}}$ exhibit all the reduction formulae together as follows:

$$p_{40}/p_{20}^2=p_{04}/p_{02}^2=3,\quad p_{31}/p_{20}^{\frac{3}{2}}p_{02}^{\frac{1}{2}}=p_{13}/p_{20}^{\frac{1}{2}}p_{02}^{\frac{3}{2}}=3\rho,\quad p_{22}/p_{20}p_{02}=1+2\rho^2,$$
$$p_{60}/p_{20}^3=p_{06}/p_{02}^3=15,\quad p_{51}/p_{20}^{\frac{5}{2}}p_{02}^{\frac{1}{2}}=p_{15}/p_{20}^{\frac{1}{2}}p_{02}^{\frac{5}{2}}=15\rho,$$
$$p_{42}/p_{20}^2p_{02}=p_{24}/p_{20}p_{02}^2=3+12\rho^2,\quad p_{33}/p_{20}^{\frac{3}{2}}p_{02}^{\frac{3}{2}}=9\rho+6\rho^3 \qquad \text{...............(xxxiii).}$$

If in like manner the numerator and denominator of the expression (xvii) for the deviation in r in terms of the deviations in the moments be divided by $\sqrt{p_{20}}\cdot\sqrt{p_{02}}$ and we write

$$\alpha_1=\frac{dp_{10}}{\sqrt{p_{20}}},\quad \alpha_2=\frac{dp_{20}}{p_{20}},\quad \beta_1=\frac{dp_{01}}{\sqrt{p_{02}}},\quad \beta_2=\frac{dp_{02}}{p_{02}},\quad \gamma=\frac{dp_{11}}{\sqrt{p_{20}}\sqrt{p_{02}}}$$

it becomes

$$\bar{r}+dr=\frac{\rho+\gamma-\alpha_1\beta_1}{\sqrt{(1+\alpha_2-\alpha_1^2)}\times\sqrt{(1+\beta_2-\beta_1^2)}} \qquad \text{............(xxxiv).}$$

When this is expanded, the mean values of α_1, β_1, α_2, β_2, and γ are of course zero. In the following tables the general formulae (xxviii), (xxix) and (xxxi) for the mean products two, three and four together of deviations of moments in samples of n are given at the head, the suffixes used and their composition in the several formulae are shown in the initial columns (omitting repetitions of dp and p to abbreviate the printed matter), and the resulting formulae for the mean products of the α's, β's and γ's required in the expansion of (xxxiv) are given in the last column, the reductions of the higher moments having been made by (xxxiii) to suit the case we are investigating, that of a normal distribution of the two characters in the material sampled. Since the first and second moments only of r are at present sought it is unnecessary to take products involving higher powers of γ than the second. The four additional formulae to γ^4 are inserted, however, to complete the formulae for the third and fourth moments when required.

As an illustration, if in equation (xxix) we put

$$u = 0, \quad v = 2, \quad u' = 1, \quad v' = 1, \quad u'' = 1, \quad v'' = 1,$$

we get $$\text{mean } dp_{02}dp_{11}dp_{11} = \frac{1}{n^2}[p_{24} - p_{13}p_{11} - p_{13}p_{11} - p_{22}p_{02} + 2p_{02}p_{11}p_{11}].$$

Dividing by $p_{20}p_{02}^2$ and using (xxxiii) it follows that

$$\text{mean } \beta_2\gamma^2 = \frac{1}{n^2}[(3 + 12\rho^2) - 3\rho^2 - 3\rho^3 - (1 + 2\rho^2) + 2\rho^2]$$

$$= \frac{1}{n^2}[2 + 6\rho^2],$$

the penultimate formula in table (xxxvi). It will be seen that the values to be attributed to the terms signified by the suffixes between the double rules are the right-hand sides of (xxxiii), the composition of the terms being shown in the last column of the formula. For 24 put $3 + 12\rho^2$, for 13 put 3ρ, for 11 put ρ, for 22 put $1 + 2\rho^2$, for 02 put 1, and the formula for mean $\beta_2\gamma^2$ may be written down without any division being necessary.

The formulae (xxxvii) are derived from those in (xxxv) and the suffixes between the double rules are for reference to the first columns of the latter table. Thus in the penultimate formula look up 02 11 in (xxxv) and find 2ρ. Look up 11 11 and find $1 + \rho^2$. $2\rho \times (1 + \rho^2) = 2\rho + 2\rho^3$, the first of the three component terms added together in the last column of the formula.

$$\text{mean } dp_{uv} \,.\, dp_{u'v'} = \frac{1}{n}\left[p_{u+u'\,v+v'} - p_{uv}p_{u'v'}\right]$$

			–		Mean value of	$=\frac{1}{n}\times$
uv	$u'v'$	$u+u'$ $v+v'$	uv	$u'v'$		
10	10	20	10	10	α_1^2	1
01	01	02	01	01	β_1^2	1
10	20	30	10	20	$\alpha_1\alpha_2$	0
01	02	03	01	02	$\beta_1\beta_2$	0
20	20	40	20	20	α_2^2	2
02	02	04	02	02	β_2^2	2
10	01	11	10	01	$\alpha_1\beta_1$	ρ
20	02	22	20	02	$\alpha_2\beta_2$	$2\rho^2$
10	02	12	10	02	$\alpha_1\beta_2$	0
20	01	21	20	01	$\alpha_2\beta_1$	0
10	11	21	10	11	$\alpha_1\gamma$	0
01	11	12	01	11	$\beta_1\gamma$	0
20	11	31	20	11	$\alpha_2\gamma$	2ρ
02	11	13	02	11	$\beta_2\gamma$	2ρ
11	11	22	11	11	γ^2	$1+\rho^2$

......(xxxv).

mean $dp_{uv} . dp_{u'v'} . dp_{u''v''}$

$$= \frac{1}{n^2}\left[p_{u+u'+u''\,v+v'+v''} - p_{u+u'\,v+v'}\,p_{u''v''} - p_{u+u''\,v+v''}\,p_{u'v'} - p_{u'+u''\,v'+v''}\,p_{uv} + 2p_{uv}\,p_{u'v'}\,p_{u''v''}\right]$$

				–		–		–		+2 ×			Mean value of	$=\frac{1}{n^2}\times$
uv	$u'v'$	$u''v''$	$u+u'+u''$ $v+v'+v''$	$u+u'$ $v+v'$	$u''v''$	$u+u''$ $v+v''$	$u'v'$	$u'+u''$ $v'+v''$	uv	uv	$u'v'$	$u''v''$		
10	10	20	40	20	20	30	10	30	10	10	10	20	$\alpha_1^2\alpha_2$	2
01	01	02	04	02	02	03	01	03	01	01	01	02	$\beta_1^2\beta_2$	2
20	20	20	60	40	20	40	20	40	20	20	20	20	α_2^3	8
02	02	02	06	04	02	04	02	04	02	02	02	02	β_2^3	8
10	20	01	31	30	01	11	20	21	10	10	20	01	$\alpha_1\alpha_2\beta_1$	2ρ
10	01	02	13	11	02	12	01	03	10	10	01	02	$\alpha_1\beta_1\beta_2$	2ρ
10	10	02	22	20	02	12	10	12	10	10	10	02	$\alpha_1^2\beta_2$	$2\rho^2$
20	01	01	22	21	01	21	01	02	20	20	01	01	$\alpha_2\beta_1^2$	$2\rho^2$
20	20	02	42	40	02	22	20	22	20	20	20	02	$\alpha_2^2\beta_2$	$8\rho^2$
20	02	02	24	22	02	22	02	04	20	20	02	02	$\alpha_2\beta_2^2$	$8\rho^2$
10	10	11	31	20	11	21	10	21	10	10	10	11	$\alpha_1^2\gamma$	2ρ
01	01	11	13	02	11	12	01	12	01	01	01	11	$\beta_1^2\gamma$	2ρ
10	01	11	22	11	11	21	01	12	10	10	01	11	$\alpha_1\beta_1\gamma$	$1+\rho^2$
20	02	11	33	22	11	31	02	13	20	20	02	11	$\alpha_2\beta_2\gamma$	$4\rho+4\rho^3$
20	20	11	51	40	11	31	20	31	20	20	20	11	$\alpha_2^2\gamma$	8ρ
02	02	11	15	04	11	13	02	13	02	02	02	11	$\beta_2^2\gamma$	8ρ
20	11	11	42	31	11	31	11	22	20	20	11	11	$\alpha_2\gamma^2$	$2+6\rho^2$
02	11	11	24	13	11	13	11	22	02	02	11	11	$\beta_2\gamma^2$	$2+6\rho^2$
11	11	11	33	22	11	22	11	22	11	11	11	11	γ^3	$6\rho+2\rho^3$

............(xxxvi).

$$
\begin{aligned}
&\text{mean } dp_{uv}.dp_{u'v'}.dp_{u''v''}.dp_{u'''v'''}\\
&\quad = \text{mean } dp_{uv}.dp_{u'v'} \times \text{mean } dp_{u''v''}.dp_{u'''v'''}\\
&\qquad + \text{mean } dp_{uv}.dp_{u''v''} \times \text{mean } dp_{u'v'}.dp_{u'''v'''}\\
&\qquad\quad + \text{mean } dp_{uv}.dp_{u'''v'''} \times \text{mean } dp_{u'v'}.dp_{u''v''}.
\end{aligned}
$$

						×		+		×		+		×		Mean value of	$=\frac{1}{n^2}\times$
uv	$u'v'$	$u''v''$	$u'''v'''$	uv	$u'v'$	$u''v''$	$u'''v'''$	uv	$u''v''$	$u'v'$	$u'''v'''$	uv	$u'''v'''$	$u'v'$	$u''v''$		
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	α_1^4	3
01	01	01	01	01	01	01	01	01	01	01	01	01	01	01	01	β_1^4	3
10	10	20	20	10	10	20	20	10	20	10	20	10	20	10	20	$\alpha_1^2\alpha_2^2$	2
01	01	02	02	01	01	02	02	01	02	01	02	01	02	01	02	$\beta_1^2\beta_2^2$	2
20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	α_2^4	12
02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	β_2^4	12
10	10	10	01	10	10	10	01	10	10	10	01	10	01	10	10	$\alpha_1^3\beta_1$	3ρ
10	01	01	01	10	01	01	01	10	01	01	01	10	01	01	01	$\alpha_1\beta_1^3$	3ρ
10	10	01	01	10	10	01	01	10	01	10	01	10	01	10	01	$\alpha_1^2\beta_1^2$	$1+2\rho^2$
10	10	02	02	10	10	02	02	10	02	10	02	10	02	10	02	$\alpha_1^2\beta_2^2$	2
20	20	01	01	20	20	01	01	20	01	20	01	20	01	20	01	$\alpha_2^2\beta_1^2$	2
20	20	20	02	20	20	20	02	20	20	20	02	20	02	20	20	$\alpha_2^3\beta_2$	$12\rho^2$
20	02	02	02	20	02	02	02	20	02	02	02	20	02	02	02	$\alpha_2\beta_2^3$	$12\rho^2$
20	20	02	02	20	20	02	02	20	02	20	02	20	02	20	02	$\alpha_2^2\beta_2^2$	$4+8\rho^4$
10	20	20	01	10	20	20	01	10	20	20	01	10	01	20	20	$\alpha_1\alpha_2^2\beta_1$	2ρ
10	01	02	02	10	01	02	02	10	02	01	02	10	02	01	02	$\alpha_1\beta_1\beta_2^2$	2ρ
10	10	20	02	10	10	20	02	10	20	10	02	10	02	10	20	$\alpha_1^2\alpha_2\beta_2$	$2\rho^2$
20	01	01	02	20	01	01	02	20	01	01	02	20	02	01	01	$\alpha_2\beta_1^2\beta_2$	$2\rho^2$
10	01	20	02	10	01	20	02	10	20	01	02	10	02	01	20	$\alpha_1\beta_1\alpha_2\beta_2$	$2\rho^3$
10	10	20	11	10	10	20	11	10	20	10	11	10	11	10	20	$\alpha_1^2\alpha_2\gamma$	2ρ
01	01	02	11	01	01	02	11	01	02	01	11	01	11	01	02	$\beta_1^2\beta_2\gamma$	2ρ
20	20	20	11	20	20	20	11	20	20	20	11	20	11	20	20	$\alpha_2^3\gamma$	12ρ
02	02	02	11	02	02	02	11	02	02	02	11	02	11	02	02	$\beta_2^3\gamma$	12ρ
10	20	01	11	10	20	01	11	10	01	20	11	10	11	20	01	$\alpha_1\alpha_2\beta_1\gamma$	$2\rho^2$
10	01	02	11	10	01	02	11	10	02	01	11	10	11	01	02	$\alpha_1\beta_1\beta_2\gamma$	$2\rho^2$
10	10	02	11	10	10	02	11	10	02	10	11	10	11	10	02	$\alpha_1^2\beta_2\gamma$	2ρ
20	01	01	11	20	01	01	11	20	01	01	11	20	11	01	01	$\alpha_2\beta_1^2\gamma$	2ρ
20	20	02	11	20	20	02	11	20	02	20	11	20	11	20	02	$\alpha_2^2\beta_2\gamma$	$4\rho+8\rho^3$
20	02	02	11	20	02	02	11	20	02	02	11	20	11	02	02	$\alpha_2\beta_2^2\gamma$	$4\rho+8\rho^3$
10	10	11	11	10	10	11	11	10	11	10	11	10	11	10	11	$\alpha_1^2\gamma^2$	$1+\rho^2$
01	01	11	11	01	01	11	11	01	11	01	11	01	11	01	11	$\beta_1^2\gamma^2$	$1+\rho^2$
10	01	11	11	10	01	11	11	10	11	01	11	10	11	01	11	$\alpha_1\beta_1\gamma^2$	$\rho+\rho^3$
20	02	11	11	20	02	11	11	20	11	02	11	20	11	02	11	$\alpha_2\beta_2\gamma^2$	$10\rho^2+2\rho^4$
20	20	11	11	20	20	11	11	20	11	20	11	20	11	20	11	$\alpha_2^2\gamma^2$	$2+10\rho^2$
02	02	11	11	02	02	11	11	02	11	02	11	02	11	02	11	$\beta_2^2\gamma^2$	$2+10\rho^2$
20	11	11	11	20	11	11	11	20	11	11	11	20	11	11	11	$\alpha_2\gamma^3$	$6\rho+6\rho^3$
02	11	11	11	02	11	11	11	02	11	11	11	02	11	11	11	$\beta_2\gamma^3$	$6\rho+6\rho^3$
11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	γ^4	$2+4\rho^2+2\rho^4$

..................(xxxvii).

We are now in a position to expand (xxxiv) to terms of the fourth degree in α, β and γ, take mean values of the α, β, γ products as found for samples of normal distributions and so obtain $\bar{r}$ the mean value of the correlation coefficient in samples of n of such distributions correct to terms in $1/n^2$.

$$\begin{aligned}
&\bar{r}+dr\\
&=(\rho+\gamma-\alpha_1\beta_1)\\
&\quad\times(1-\tfrac{1}{2}\alpha_2+\tfrac{1}{2}\alpha_1^2+\tfrac{3}{8}\alpha_2^2-\tfrac{3}{4}\alpha_1^2\alpha_2-\tfrac{5}{16}\alpha_2^3+\tfrac{3}{8}\alpha_1^4+\tfrac{15}{16}\alpha_1^2\alpha_2^2+\tfrac{35}{128}\alpha_2^4)\\
&\quad\times(1-\tfrac{1}{2}\beta_2+\tfrac{1}{2}\beta_1^2+\tfrac{3}{8}\beta_2^2-\tfrac{3}{4}\beta_1^2\beta_2-\tfrac{5}{16}\beta_2^3+\tfrac{3}{8}\beta_1^4+\tfrac{15}{16}\beta_1^2\beta_2^2+\tfrac{35}{128}\beta_2^4)\\
&=\rho-\tfrac{1}{2}\{\rho(\alpha_2+\beta_2)-2\gamma\}\\
&\quad+\tfrac{1}{8}\{4\rho(\alpha_1^2+\beta_1^2)+3\rho(\alpha_2^2+\beta_2^2)+2\rho\alpha_2\beta_2-4(\alpha_2\gamma+\beta_2\gamma)-8\alpha_1\beta_1\}\\
&\quad+\tfrac{1}{16}\{-12\rho(\alpha_1^2\alpha_2+\beta_1^2\beta_2)-5\rho(\alpha_2^3+\beta_2^3)-4\rho(\alpha_1^2\beta_2+\alpha_2\beta_1^2)-3\rho(\alpha_2^2\beta_2+\alpha_2\beta_2^2)\\
&\qquad\qquad+8(\alpha_1^2\gamma+\beta_1^2\gamma)+6(\alpha_2^2\gamma+\beta_2^2\gamma)+4\alpha_2\beta_2\gamma+8(\alpha_1\alpha_2\beta_1+\alpha_1\beta_1\beta_2)\}\\
&\quad+\tfrac{1}{128}\{48\rho(\alpha_1^4+\beta_1^4)+120\rho(\alpha_1^2\alpha_2^2+\beta_1^2\beta_2^2)+35\rho(\alpha_2^4+\beta_2^4)\\
&\quad+48\rho(\alpha_1^2\alpha_2\beta_2+\alpha_2\beta_1^2\beta_2)+20\rho(\alpha_2^3\beta_2+\alpha_2\beta_2^3)+32\rho\alpha_1^2\beta_1^2+24\rho(\alpha_1^2\beta_2^2+\alpha_2^2\beta_1^2)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad+18\rho\alpha_2^2\beta_2^2\\
&\quad-96(\alpha_1^2\alpha_2\gamma+\beta_1^2\beta_2\gamma)-40(\alpha_2^3\gamma+\beta_2^3\gamma)-32(\alpha_1^2\beta_2\gamma+\alpha_2\beta_1^2\gamma)-24(\alpha_2^2\beta_2\gamma+\alpha_2\beta_2^2\gamma)\\
&\quad-64(\alpha_1^3\beta_1+\alpha_1\beta_1^3)-48(\alpha_1\alpha_2^2\beta_1+\alpha_1\beta_1\beta_2^2)-32\alpha_1\beta_1\alpha_2\beta_2\}\ \ldots\ldots\ldots\ldots\ldots\text{(xxxviii)}.
\end{aligned}$$

Whence taking mean values in samples of n of a normal distribution,

$$\begin{aligned}
\bar{r}=\rho&+\frac{1}{8n}\{8\rho+12\rho+4\rho^3-16\rho-8\rho\}\\
&+\frac{1}{16n^2}\{-48\rho-80\rho-16\rho^3-48\rho^3\\
&\qquad\qquad+32\rho+96\rho+16\rho+16\rho^3+32\rho\}\\
&+\frac{1}{128n^2}\{288\rho+480\rho+840\rho\\
&\qquad\qquad+192\rho^3+480\rho^3+32\rho+64\rho^3+96\rho+72\rho+144\rho^5\\
&\qquad\qquad-384\rho-960\rho-128\rho-192\rho-384\rho^3\\
&\qquad\qquad-384\rho-192\rho-64\rho^3\}\\
=\rho&-\frac{1}{2n}\rho(1-\rho^2)-\frac{3}{8n^2}\rho(1-\rho^2)(1+3\rho^2)\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxxix)}.
\end{aligned}$$

Or, expressing the result in terms of $n-1$ (by changing n into $n'+1$ and expanding) we may write to the same degree of approximation

$$\bar{r}=\rho\left[1-\frac{1-\rho^2}{2(n-1)}\left\{1-\frac{1-9\rho^2}{4(n-1)}\right\}\right]\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xl)},$$

from which follows that

$$1-\bar{r}^2=(1-\rho^2)\left[1+\frac{\rho^2}{n-1}\left\{1-\frac{1-5\rho^2}{2(n-1)}\right\}\right]\ldots\ldots\ldots\ldots\ldots\text{(xli)}.$$

And, again, by squaring (xxxiv) and expanding we obtain in the same way

$$\bar{r}^2 + 2\bar{r}\,dr + (dr)^2$$

$$= (\rho^2 + 2\rho\gamma + \gamma^2 - 2\rho\alpha_1\beta_1 - 2\alpha_1\beta_1\gamma + \alpha_1^2\beta_1^2)$$

$$\times (1 - \alpha_2 + \alpha_1^2 + \alpha_2^2 - 2\alpha_1^2\alpha_2 - \alpha_2^3 + \alpha_1^4 + 3\alpha_1^2\alpha_2^2 + \alpha_2^4)$$

$$\times (1 - \beta_2 + \beta_1^2 + \beta_2^2 - 2\beta_1^2\beta_2 - \beta_2^3 + \beta_1^4 + 3\beta_1^2\beta_2^2 + \beta_2^4)$$

$$= \rho^2 + \{2\rho\gamma - \rho^2\alpha_2 - \rho^2\beta_2\}$$

$$+ \{\gamma^2 - 2\rho(\alpha_1\beta_1 + \alpha_2\gamma + \beta_2\gamma) + \rho^2(\alpha_1^2 + \beta_1^2 + \alpha_2^2 + \beta_2^2 + \alpha_2\beta_2)\}$$

$$+ \{-2\alpha_1\beta_1\gamma - \alpha_2\gamma^2 - \beta_2\gamma^2 + 2\rho(\alpha_1\alpha_2\beta_1 + \alpha_1\beta_1\beta_2 + \alpha_1^2\gamma + \beta_1^2\gamma + \alpha_2^2\gamma + \beta_2^2\gamma + \alpha_2\beta_2\gamma)$$

$$+ \rho^2(-2\alpha_1^2\alpha_2 - 2\beta_1^2\beta_2 - \alpha_2^3 - \beta_2^3 - \alpha_1^2\beta_2 - \alpha_2\beta_1^2 - \alpha_2^2\beta_2 - \alpha_2\beta_2^2)\}$$

$$+ \{\alpha_1^2\beta_1^2 + 2\alpha_1\alpha_2\beta_1\gamma + 2\alpha_1\beta_1\beta_2\gamma + \alpha_1^2\gamma^2 + \beta_1^2\gamma^2 + \alpha_2^2\gamma^2 + \beta_2^2\gamma^2 + \alpha_2\beta_2\gamma^2$$

$$+ 2\rho(-\alpha_1^3\beta_1 - \alpha_1\beta_1^3 - \alpha_1\alpha_2^2\beta_1 - \alpha_1\beta_1\beta_2^2 - \alpha_1\beta_1\alpha_2\beta_2$$

$$- 2\alpha_1^2\alpha_2\gamma - 2\beta_1^2\beta_2\gamma - \alpha_2^3\gamma - \beta_2^3\gamma - \alpha_1^2\beta_2\gamma - \alpha_2\beta_1^2\gamma - \alpha_2^2\beta_2\gamma - \alpha_2\beta_2^2\gamma)$$

$$+ \rho^2(\alpha_1^4 + \beta_1^4 + 3\alpha_1^2\alpha_2^2 + 3\beta_1^2\beta_2^2 + \alpha_2^4 + \beta_2^4 + 2\alpha_1^2\alpha_2\beta_2 + 2\alpha_2\beta_1^2\beta_2$$

$$+ \alpha_2^3\beta_2 + \alpha_2\beta_2^3 + \alpha_1^2\beta_1^2 + \alpha_2^2\beta_2^2 + \alpha_1^2\beta_2^2 + \alpha_2^2\beta_1^2)\} \quad \text{......................(xlii)}.$$

And taking mean values in samples of n of a normal distribution,

$$\bar{r}^2 + \text{mean}\,(dr)^2$$

$$= \rho^2 + \frac{1}{n}\{1 + \rho^2 - 2\rho(\rho + 4\rho) + \rho^2(2 + 4 + 2\rho^2)\}$$

$$+ \frac{1}{n^2}\{-2 - 2\rho^2 - 4 - 12\rho^2 + 2\rho(4\rho + 4\rho + 16\rho + 4\rho + 4\rho^3)$$

$$+ \rho^2(-8 - 16 - 4\rho^2 - 16\rho^2)\}$$

$$+ \frac{1}{n^2}\{1 + 2\rho^2 + 8\rho^2 + 2 + 2\rho^2 + 4 + 20\rho^2 + 10\rho^2 + 2\rho^4$$

$$+ 2\rho(-6\rho - 4\rho - 2\rho^3 - 8\rho - 24\rho - 4\rho - 8\rho - 16\rho^3)$$

$$+ \rho^2(6 + 12 + 24 + 8\rho^2 + 24\rho^2 + 1 + 2\rho^2 + 4 + 8\rho^4 + 4)\}$$

$$= \rho^2 + \frac{1}{n}\{1 - 3\rho^2 + 2\rho^4\}$$

$$+ \frac{1}{n^2}\{-6 + 18\rho^2 - 12\rho^4\}$$

$$+ \frac{1}{n^2}\{7 - 15\rho^2 + 8\rho^6\}$$

$$= \rho^2 + \frac{1}{n}(1 - \rho^2)(1 - 2\rho^2) + \frac{1}{n^2}(1 - \rho^2)(1 + 4\rho^2 - 8\rho^4).$$

And by squaring (xxxix),

$$\bar{r}^2 = \rho^2 - \frac{1}{n}\rho^2(1 - \rho^2) - \frac{1}{2n^2}\rho^2(1 - \rho^2)(1 + 5\rho^2).$$

Hence by subtraction,

$$\text{mean } (dr)^2 = \frac{1}{n}(1-\rho^2)^2 + \frac{1}{n^2}(1-\rho^2)^2(1+5\tfrac{1}{2}\rho^2).$$

Or taking the square root

$$\sigma_r = \frac{1-\rho^2}{\sqrt{n}}\left(1+\frac{1+5\frac{1}{2}\rho^2}{2n}\right),$$

which may be expressed in like manner with $\bar{r}$ in the form

$$\sigma_r = \frac{1-\rho^2}{\sqrt{n-1}}\left[1+\frac{11\rho^2}{4(n-1)}\right] \quad\text{..........................(xliii),}$$

to the same degree of approximation.

It appears then from the above results that if the coefficient of correlation existing between two measured characters in a large aggregate of individuals be computed from the product moment values in small samples, these values are subject to errors from a mean value, the standard deviation of which errors may be very approximately represented by the formula

$$\frac{1-\rho^2}{\sqrt{n-1}},$$

and with greater degree of accuracy by the formula

$$\frac{1-\rho^2}{\sqrt{n-1}}\left(1+\frac{11\rho^2}{4n}\right),$$

ρ being the coefficient of correlation between the characters in the material sampled and n being the number in the sample.

Moreover the mean value of the correlation coefficients obtained from such small samples will be less than the true coefficient of the aggregate and will be approximately represented by the formula

$$\rho\left(1-\frac{1-\rho^2}{2n}\right),$$

the defect being very small when ρ is large, and when ρ is small being of the order 5 °/ₒ in samples of 10 and ·5 °/ₒ in samples of 100.

On the other hand the modal value of the correlation coefficients, or the most likely value in a single sample, will be greater than the true correlation coefficient (that is to say numerically greater: the correlation being supposed to be measured positively).

We have, by definition $\quad 1/\lambda = \dfrac{\sigma_r^2}{1-\bar{r}^2},$

and so from equations (xli), (xliii) putting $n-1=n'$

$$1/\lambda = \frac{1-\rho^2}{n'}\left(1+\frac{11\rho^2}{4n'}\right)\left(1+\frac{\rho^2}{n'}\right)^{-1}$$

$$= \frac{1-\rho^2}{n'}\left(1+\frac{7\rho^2}{4n'}\right),$$

using second approximations.

And hence from (x) going now to third approximations,

$$\begin{aligned}
\breve{r} &= \bar{r}\,\frac{1-1/\lambda}{1-3/\lambda} \\
&= \rho\left\{1-\frac{1-\rho^2}{2n'}+\frac{(1-\rho^2)(1-9\rho^2)}{8n'^2}\right\} \\
&\quad\times\left\{1-\frac{1-\rho^2}{n'}-\frac{(1-\rho^2)\,9\rho^2}{2n'^2}\right\} \\
&\quad\times\left\{1-\frac{3(1-\rho^2)}{n'}-\frac{(1-\rho^2)\,27\rho^2}{2n'^2}\right\}^{-1} \\
&= \rho\left\{1+\frac{3(1-\rho^2)}{2n'}+\frac{(41+23\rho^2)(1-\rho^2)}{8n'^2}\right\}.
\end{aligned}$$

The excess of $\breve{r}$ over true ρ is zero if $\rho=0$ and if $\rho=1$, but if n is small and ρ fairly large the excess may be such as to make the modal value unity or greater than unity. If for instance n is so small as 4, n' being thus 3, the above approximate equation gives

$$\breve{r}=\rho+\tfrac{1}{2}\rho(1-\rho^2)+(\tfrac{41}{72}\rho+\tfrac{23}{72}\rho^3)(1-\rho^2)$$

or

$$\begin{aligned}
\breve{r} &= 2{\cdot}069\rho-{\cdot}750\rho^3-{\cdot}319\rho^5 \\
&= {\cdot}93 \text{ when } \rho={\cdot}5 \\
&= 1{\cdot}05 \quad ,, \quad \rho={\cdot}6 \\
&= 1{\cdot}14 \quad ,, \quad \rho={\cdot}7.
\end{aligned}$$

The frequency distribution in the last two cases is of the J type, there being no mode within the range. The greatest frequency is at the extremity of the range, or at value unity. The interpretation of this result is clearly that such small samples as 3, 4 or 5, as might be expected, fail altogether to give by the product moment formula an approximation to the correlation coefficient. Under some circumstances the points which graphically represent the observed measures are more likely to be in a line than to have a configuration represented by any specified fractional correlation coefficient. This will happen if the correlation in the material has a larger coefficient than ·6 (approx.) when samples of four are drawn: or a larger coefficient than ·3 (approx.) when samples of three are drawn. If samples of two are drawn the coefficient of correlation is necessarily unity in the sample whatever it may be in the material*. All the distribution is concentrated at value unity and $\breve{r}$ should in this case be infinite for all values of ρ. Our approximation, neglecting terms in $1/n'^3$ etc., cannot of course show this if $n'=1$. It gives $\breve{r}$ greater than unity and so a J type, but fails to show the complete concentration at unity.

It appears from (x) that $\breve{r}$ will be infinite when $\lambda=3$ and $\bar{r}$ any value other than zero, whilst $\breve{r}$ will be zero if $\bar{r}$ is zero and λ other than 3. If $\bar{r}$, and therefore

* Supposing the material ungrouped. If it is grouped some values will be indeterminate in small samples, viz. when all observations fall into the same group.

ρ, is zero and $\lambda = 3$, $\check{r}$ is indeterminate. This case is a little singular and (ix) shows that m_1 and m_2 are zero and that the frequency is therefore constant or that all values of r are equally likely, from -1 to $+1$. Remembering that $\lambda = \frac{1-\bar{r}^2}{\sigma_r^2}$ it will happen when $\sigma_r^2 = \frac{1}{3}$. But when $\rho = 0$ the value we have found for σ_r^2 is $\frac{1}{n-1}$ which will be $\frac{1}{3}$ when $n = 4$. If therefore from material possessing zero correlation samples of four are drawn, the values of the correlation coefficients in these samples should, if the above formulae and assumed type of frequency distribution are correct, be equally distributed in value throughout the whole range.

Complete experimental confirmation of the distributions found in this paper for the product moment expression of the coefficient of correlation in small samples is difficult to obtain. The second order differences in the values of the mean and standard deviation of the distribution are necessarily small in large samples and comparable with the errors of sampling of such samples unless a very large *number* of samples is taken. On the other hand if small samples are taken the theory is not so exact and in addition the distribution tends to concentrate in certain grades if the original material is coarsely grouped, whilst fine grouping adds very much to the labour of computation.

In a paper published in *Biometrika*, Vol. VI., p. 302, "Student" gives the results of a very painstaking investigation to determine experimentally these distributions. Material based on W. R. Macdonell's measures of finger length and stature in 3000 criminals [*Biometrika*, Vol. I., p. 219] was formed having correlation with coefficient* ·66 and samples of 4, 8 and 30 were drawn. Material with correlation zero was also constructed and samples of 4 and 8 drawn.

The comparison of the actual means and standard deviations found in these experiments with those computed from the formulae in this paper is shown in the table that follows.

	1	*2*	*3*	*4*	*5*
Number in sample, n	4	8	4	8	30
Correlation coefficient in material, ρ	0	0	·66	·66	·66
Mean correlation coefficient in samples, observed ...	—	—	·5609	·6139	·661
" " " " calculated, $\bar{r}$ ...	0	0	·5933	·6317	·6542
Standard deviation of coefficient in samples, observed ...	·5512	·3731	·4680	·2684	·1001
" " " " calculated, σ_r	·5773	·3780	·4234	·2453	·1088
Number of samples taken	745	750	745	750	100
Probable error of mean	·0143	·0093	·0073	·0105	·0060
" " standard deviation	·0101	·0066	·0052	·0074	·0043

* The exact figure was ·6608.

It will be seen that the differences of observed and calculated values are for the most part several times the probable errors. Estimated in this way case (3) $\rho = \cdot 66$, $n = 4$, is the worst fit, the difference of the mean being five times and that of the standard deviation nine times the probable errors.

From the values of $\bar{r}$ and σ_r found above are calculated

	1	*2*	*3*	*4*	*5*
$\lambda = \dfrac{1-\bar{r}^2}{\sigma_r^2} =$	3	7	3·6140	9·9905	48·325
$\breve{r} = \dfrac{\lambda - 1}{\lambda - 3} \cdot \bar{r} =$	$\frac{0}{0}$	0	2·5259	·8124	·6831

The frequency distributions of Type I to fit the numbers of samples taken in the experiments and the values of $\bar{r}$ and σ_r calculated will be of the form

$$y = y_0 (1 - x)^{m_1} (1 + x)^{m_2}$$

when referred to the absolute origin and unit of measurement of r, and the constants* will be

	1	*2*	*3*	*4*	*5*
$m_1 = \frac{1}{2}(\lambda - 1)(1 - \breve{r}) - 1 =$	0	2	− ·46844	·6557	7·1825
$m_2 = \frac{1}{2}(\lambda - 1)(1 + \breve{r}) - 1 =$	0	2	1·08243	6·3348	38·1425
$y_0 = \dfrac{N}{2^{m_1+m_2+1}} \cdot \dfrac{\Gamma(m_1+m_2+2)}{\Gamma(m_1+1) \cdot \Gamma(m_2+1)} =$	372·5	703·12	202·25	95·131	·0033889

When the above frequency curves are plotted they appear as shown on the diagrams pp. 112, 113 and are seen to be in fair consonance with the frequencies observed in the experiments and shown by the rectangles upon the same diagrams. They are perhaps as good an expression of these frequencies as could be found amongst the type of frequency curve assumed. The case of $\rho = 0$, $n = 4$, for which theory prescribes a horizontal straight line is seen to be very nearly so in the experiment, apart from individual fluctuations. In $\rho = 0$, $n = 8$, the curve well fits the deviations from zero correlation observed. In $\rho = \cdot 66$, $n = 4$, the asymptotic nature of the distribution towards the value unity which the fitted curve foreshadows is borne out in the samples of four drawn. With larger samples from the same class of material the displacement of the mode and the skewness of the distribution resulting from the assumed types are corroborated in the tests.

At the same time it must be admitted that there are considerable differences

* For the special case of $\rho = 0$, eqns. (ix), (xl) and (xliii) show that the curve is $y = y_0 (1 - x^2)^{\frac{1}{2}(n-4)}$, the form suggested for this case by "Student," *Biometrika*, Vol. VI. p. 306.

needing to be accounted for. The individual irregularities in the observations are more than would be expected in random samples of homogeneous material and it is probable that these jumps which make a good fit of any continuous curve whatever an impossibility are partly due to grouping. If the grouping of the original material were too coarse there would be a tendency in small samples for statistical constants to centre round certain values. Another possible source of error is in the mixing and in the drawing out of samples. Although a great deal of trouble was undoubtedly taken in these experiments, yet there always seems room for a little involuntary order in repetitions intended to go solely by chance.

The curves were planimetered and the following tables show the comparison of theoretical with observed frequencies of the grade values. It will be seen that the differences are not systematic but that the + and − errors are fairly mixed. On calculating the square contingency, χ^2, and deducing from this and the number of groups, n', the probability, P, of these differences being purely that of sampling* such probability comes out in most cases very small. It seems legitimate in this instance to see to what extent grouping will smooth down the irregularities and yet show the general resemblance and, with this in view, the differences in the columns headed e' are calculated and the grouping therein shown may be taken to indicate what is necessary to bring the probabilities within reasonable distance of expectation.

	$\rho = 0, \quad n = 4.$				$\rho = 0, \quad n = 8.$			
r	Calculated frequency m	Observed frequency	Difference e	$\frac{e^2}{m}$	Calculated frequency m	Observed frequency	Difference e	$\frac{e^2}{m}$
·925—1·0	27·94	22·5	− 5·44	1·06	·6	—	− ·6	·60
·825—	37·25	31·5	− 5·75	·89	4·1	3	− 1·1	·30
·725—	37·25	24·0	−13·25	4·71	11·6	12	+ ·4	·01
·625—	37·25	35·0	− 2·25	·14	20·6	11·5	− 9·1	4·02
·525—	”	34·0	− 3·25	·28	31·9	28·5	− 3·4	·36
·425—	”	47·0	+ 9·75	2·55	42·3	46	+ 3·7	·32
·325—	”	30·5	− 6·75	1·22	51·6	47·5	− 4·1	·33
·225—	”	46·5	+ 9·25	2·30	59·9	70	+10·1	1·70
·125—	”	44·0	+ 6·75	1·22	65·8	57·5	− 8·3	1·05
·025—	”	32·0	− 5·25	·74	69·4	70	+ ·6	·01
$\bar{1}$·925—	”	45·0	+ 7·75	1·61	70·3	60·5	− 9·8	1·37
$\bar{1}$·825—	”	43·0	+ 5·75	·89	67·9	71·5	+ 3·6	·19
$\bar{1}$·725—	”	41·0	+ 3·75	·38	63·1	76	+12·9	2·64
$\bar{1}$·625—	”	37·0	− ·25	·00	56·3	63	+ 6·7	·80
$\bar{1}$·525—	”	44·0	+ 6·75	1·22	47·1	42	− 5·1	·55
$\bar{1}$·425—	”	40·0	+ 2·75	·20	36·9	33	− 3·9	·41
$\bar{1}$·325—	”	38·5	+ 1·25	·04	26·1	29	+ 2·9	·32
$\bar{1}$·225—	”	32·5	− 4·75	·61	15·3	20	+ 4·7	1·44
$\bar{1}$·125—	”	36·0	− 1·25	·04	7·4	8	+ ·6	·05
$\bar{1}$—$\bar{1}$·125	46·56	41·0	− 5·56	·83	1·8	1	− ·8	·36
		745		20·93		750		16·83

$n' = 20, \quad \chi^2 = 20·93, \quad P = ·341.$ $\qquad$ $n' = 20, \quad \chi^2 = 16·83, \quad P = ·601.$

* Tables for testing the goodness of fit, W. Palin Elderton, *Biometrika*, Vol. I. p. 155.

Fig. 1.

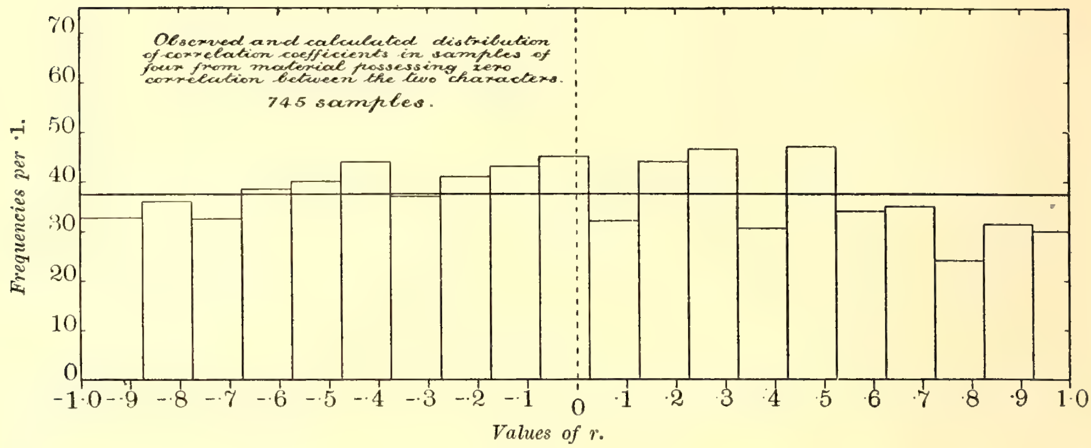

Observed and calculated distribution of correlation coefficients in samples of four from material possessing zero correlation between the two characters.
745 samples.
Frequencies per ·1.
Values of r.

Fig. 3.

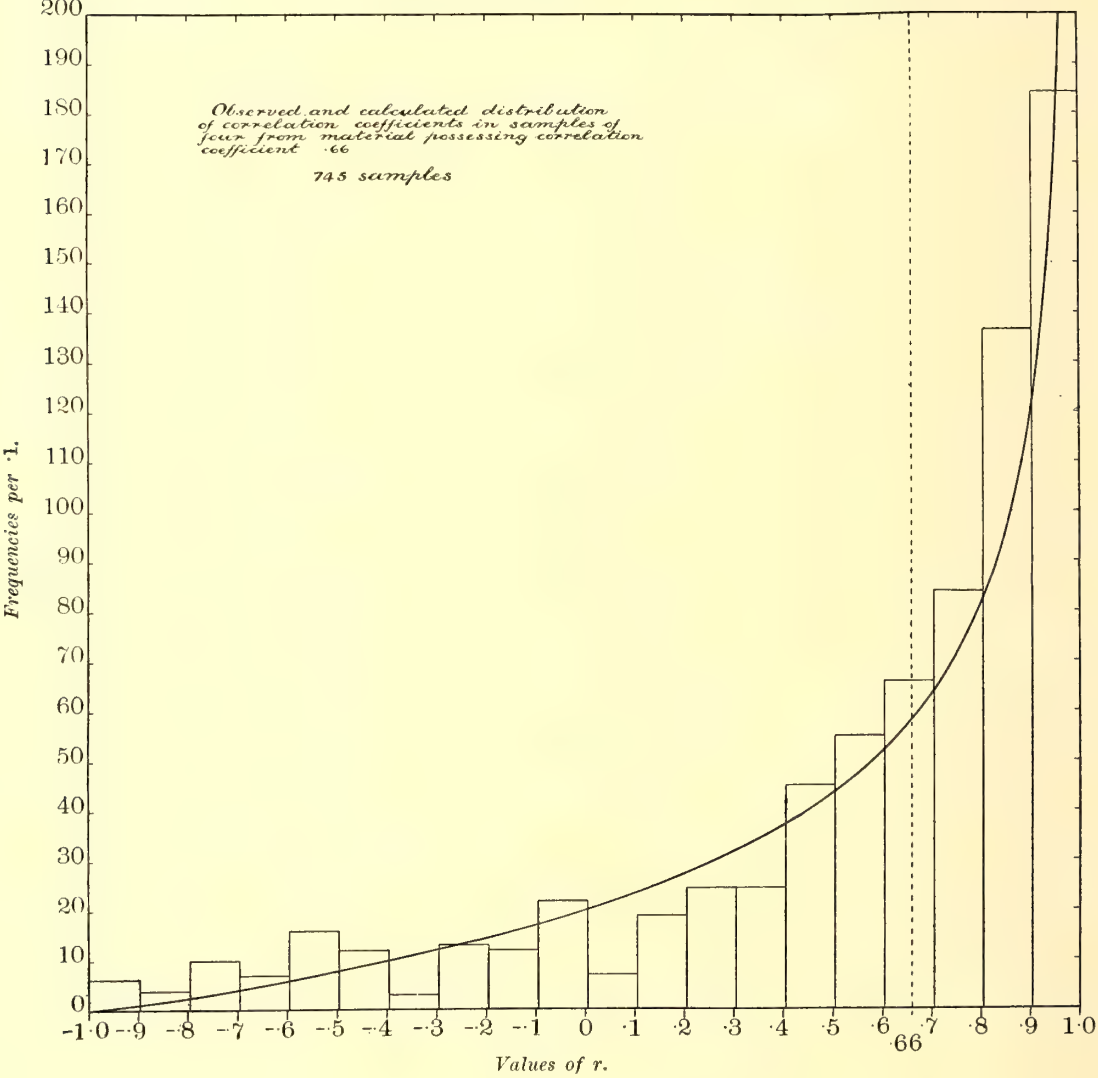
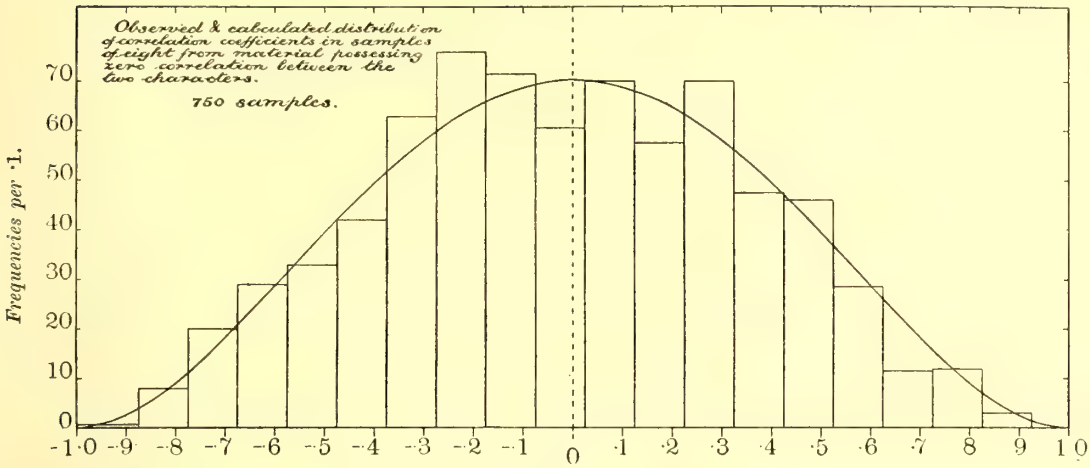
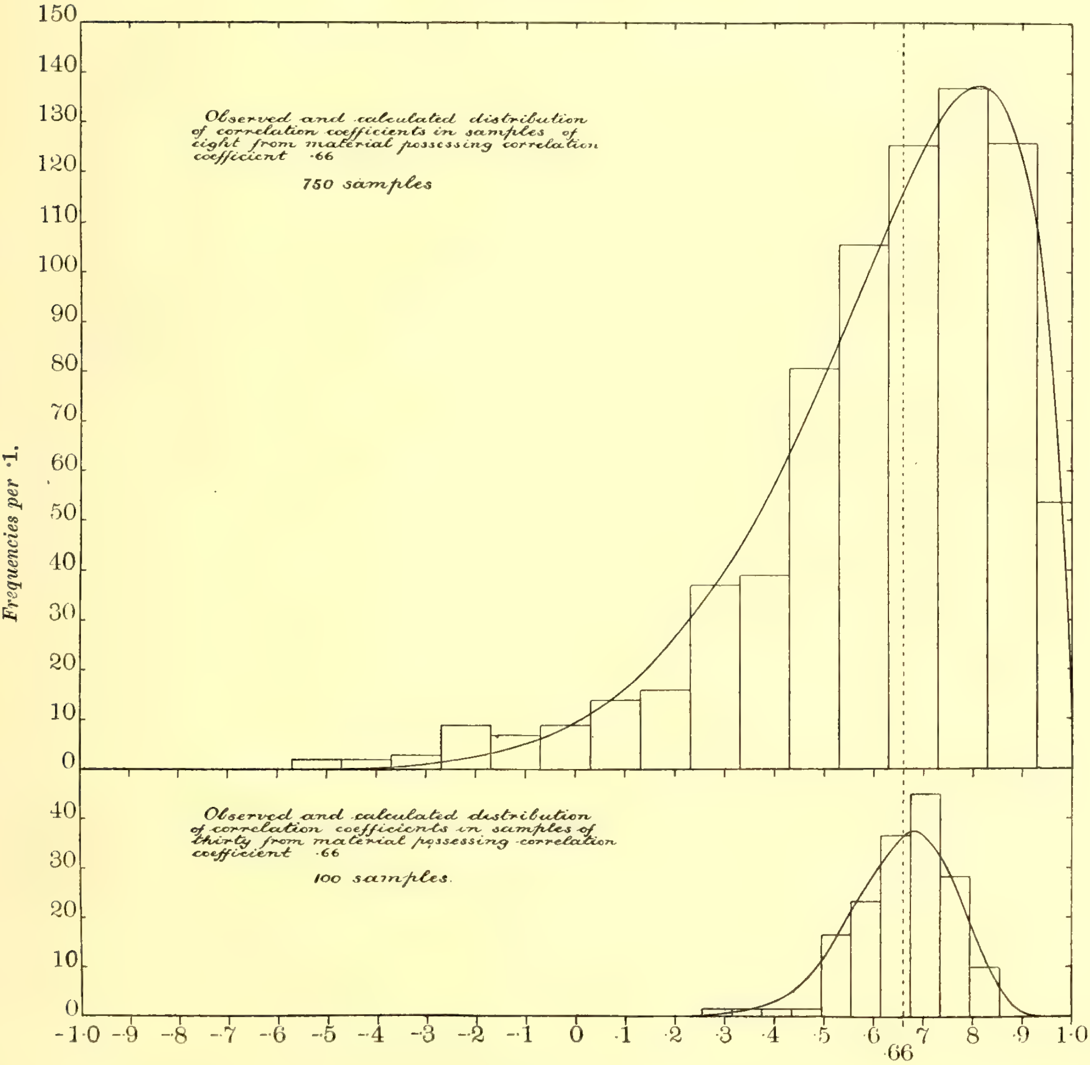

Observed and calculated distribution of correlation coefficients in samples of four from material possessing correlation coefficient ·66
745 samples
Frequencies per ·1.
·66
Values of r.

Fig. 2.



Figs. 4 and 5.



$\rho = \cdot 66, \quad n = 4.$

r	Calculated frequency m	Observed frequency	Difference e	$\frac{e^2}{m}$	Difference e'	$\frac{e'^2}{m}$
·905—1·0	230·3	175·5	− 17·2	·90	− 17·2	·90
·805—	98·9	136·5				
·705—	72·1	84	+ 20·3	3·18	+ 20·3	3·18
·605—	57·6	66				
·505—	48·0	55	+ 11·8	1·58	+ 2·0	·03
·405—	40·2	45				
·305—	34·3	24·5	− 15·0	3·52		
·205—	29·7	24·5				
·105—	25·6	19	− 21·6	9·80	− 5·1	·16
·005—	22·0	7				
$\bar{1}$·905—	18·8	22	− ·8	·02		
$\bar{1}$·805—	16·0	12				
$\bar{1}$·705—	13·5	13	− 8·7	3·06		
$\bar{1}$·605—	11·2	3				
$\bar{1}$·505—	9·0	12	+ 12·1	9·21		
$\bar{1}$·405—	6·9	16				
$\bar{1}$·305—	5·1	7	+ 8·6	8·80		
$\bar{1}$·205—	3·3	10				
$\bar{1}$·105—	1·9	4	+ 10·5	44·10		
$\bar{1}$—$\bar{1}$·105	·6	9				
	745			84·17		4·27

$n' = 10, \quad \chi^2 = 84{\cdot}17, \quad P$ very small.

$n' = 4, \quad \chi^2 = 4{\cdot}27, \quad P = {\cdot}237.$

$\rho = \cdot 66, \quad n = 8.$

r	Calculated frequency m	Observed frequency	Difference e	$\frac{e^2}{m}$	Difference e'	$\frac{e'^2}{m}$
·925—1·0	48·9	37·5	− 11·4	2·67	− 13·2	·99
·825—	127·6	126	− 1·8	·02		
·725—	135·25	137	+ 1·6	·02	+ 1·6	·02
·625—	120·0	125·5	+ 5·4	·24	+ 5·4	·24
·525—	97·25	105·5	+ 8·1	·68	+ 8·1	·68
·425—	73·9	80·5	+ 6·5	·58	− 8·0	·51
·325—	53·5	39	− 14·6	3·96		
·225—	37·0	37	0·0	·00		
·125—	24·25	16	− 8·3	2·82	+ 6·2	·41
·025—	14·75	14	− ·8	·04		
$\bar{1}$·925—	8·5	9	+ ·5	·03		
$\bar{1}$·825—	4·75	7	+ 2·2	1·05		
$\bar{1}$·725—	2·25	9	+ 6·75	20·25		
$\bar{1}$·625—	·9	3	+ 2·1	5·22		
$\bar{1}$·525—	·25	2	+ 1·75	12·24		
$\bar{1}$·425—	·10	2	+ 1·9	36·10		
$\bar{1}$·325—						
$\bar{1}$·225—						
$\bar{1}$·125—						
$\bar{1}$—$\bar{1}$·125						
	750			85·92		2·85

$n' = 16, \quad \chi^2 = 85{\cdot}92, \quad P$ very small.

$n' = 6, \quad \chi^2 = 2{\cdot}85, \quad P = {\cdot}722.$

$\rho = \cdot 66, \quad n = 30.$

r	Calculated frequency m	Observed frequency	Difference e	$\frac{e^2}{m}$
·795—	8·5	6	$-2\cdot 5$	·74
·735—	15·8	17	$+1\cdot 2$	·09
·675—	21·6	27	$+5\cdot 4$	1·35
·615—	20·9	22	$+1\cdot 1$	·06
·555—	15·8	14	$-1\cdot 8$	·21
·495—	9·8	10	$+\ \cdot 2$	·04
$\bar{1}$—·495	7·6	4	$-3\cdot 6$	1·71
	100			4·20

$n' = 7, \quad \chi^2 = 4\cdot 20, \quad P = \cdot 650.$

It is hoped that further experiments may be shortly carried out which will have regard to the points raised and show definitely whether the distributions theoretically arrived at in this paper are good presentations of fact and whether the application of the standard types of frequency curves to the distributions of statistical constants in small samples is justified.

I am indebted to Professor Pearson for drafting the lines of this investigation and for critical supervision.

ON THE MEASUREMENT OF THE INFLUENCE OF "BROAD CATEGORIES" ON CORRELATION.

By KARL PEARSON, F.R.S.

(1) By a "broad category" I understand one of a finite small number of groups into which we class a variable. For example: we may divide General Health into the categories Very Robust, Robust, Normally Healthy, Rather Delicate, Delicate and Very Delicate. These categories may be verbally defined or have their boundaries determined by quantitative limits as when we state that the limits of the Delicate coincide with so many weeks of sickness or of absence from work in the year. Again we may put into four to six classes the competitors in an examination, and the boundaries to these classes may be really percentages of marks gained. Such broad categories are very common not only in social investigations, but also in psychological records, and quite recently Dr G. A. Jaederholm, a Swedish psychologist, wrote to me asking what was the correlation between the true quantitative value of a variate in any individual and that individual's category or class-mark. The answer is an obvious one, but I do not know that I have seen it stated or any discussion of it given. It of course assumes that at the back of the categorical classification a true quantitative value lies.

Suppose a population of N individuals divided into p classes, and that C_s is the class-mark of an individual in the sth category, whose true variate is x. The problem is what is the correlation of x and the class-mark. Let $\bar{x}_s$ be the mean variate of the group of n_s individuals who fall into the sth class. Then it is just as reasonable to call $\bar{x}_s$ the class-mark as C_s, for given one the other is fixed. We really want then the correlation of x and $\bar{x}_s$.

We may either find this directly* or indirectly, and the latter is the easier course. Clearly if $\tilde{x}$ be the mean value of x for a given class, then

$$\tilde{x} = \bar{x}_s,$$

or

$$\tilde{x} - \bar{x} = \bar{x}_s - \bar{x},$$

* Let x within the class $= \bar{x}_s + x_s'$, $S =$ sum for classes and $\Sigma =$ sum within class. Then

$$S\Sigma\,(n_x x \bar{x}_s) = S\Sigma\,\{n_x \bar{x}_s (\bar{x}_s + x_s')\} = S\,(n_s \bar{x}_s^2) + S\,\{\bar{x}_s \Sigma\,(n_x x_s')\},$$

but

$$\Sigma\,(n_x x_s') = 0,$$

$$\therefore\ S\Sigma\,(n_x x \bar{x}_s) = S\,(n_s \bar{x}_s^2),$$

$$r_{x\bar{x}_s} = \frac{S\Sigma\,(n_x x \bar{x}_s)}{N\,.\,\sigma_{\bar{x}_s} \times \sigma_x} = \frac{S\,(n_s \bar{x}_s^2)}{N \sigma_{\bar{x}_s} \sigma_x} = \frac{\sigma^2_{\bar{x}_s}}{\sigma_{\bar{x}_s} \sigma_x}$$

$$= \sigma_{\bar{x}_s}/\sigma_x \text{ as before.}$$

is the equation to the regression line; therefore

$$\frac{\tilde{x}-\bar{x}}{\sigma_x}=\frac{\sigma_{\bar{x}_s}}{\sigma_x}\times\frac{\bar{x}_s-\bar{x}}{\sigma_{\bar{x}_s}}.$$

Whence it follows that

$$r_{xC_s}=\frac{\sigma_{\bar{x}_s}}{\sigma_x}\text{ ....(i)},$$

or the correlation of a variate with its class-mark is the ratio of the standard deviation of the means of class-marks to the standard deviation of the variate.

It follows from this that in classifying a variable into broad categories the test of their efficiency, as far as number and arrangement are concerned, lies in the standard deviation of their means not differing widely from the standard deviation of the variate itself.

(2) Clearly
$$\sigma^2_{\bar{x}_s}=S\,(n_s\bar{x}_s^2)/N\text{ ....(ii)},$$
where S is a sum involving all classes.

If before classifying into broad categories we have made a quantitative determination of these classes on a sample frequency, the values of $\bar{x}_s$'s can be determined. If this has not been done, or cannot be done, we are bound to assume some form of frequency distribution. Suppose we take a normal distribution, then we know that

$$n_s\bar{x}_s=\frac{N}{\sqrt{2\pi}\,\sigma_x}\int_{x_{s_1}}^{x_{s_2}}xe^{-\frac{1}{2}\frac{x^2}{\sigma^2}}\,dx=N\sigma_x\,(z_{s_1}-z_{s_2}),$$

where
$$z=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x/\sigma_x)^2},$$

and can be found from Sheppard's Tables of the ordinates of the normal curve as soon as n_s, etc. are known, for the z's are the ordinates at start and finish of the sth class, reduced by the factor N/σ_x.

Hence

$$\sigma^2_{\bar{x}_s}=S\left\{\frac{N}{n_s}(z_{s_1}-z_{s_2})^2\times\sigma_x^2\right\},$$

or
$$r_{xC_s}=\sqrt{S\left\{\frac{N}{n_s}(z_{s_1}-z_{s_2})^2\right\}}\text{ ....(iii)}.$$

Thus r_{xC_s} can be found at once from Sheppard's Tables, when the totals of the broad classes are known.

(3) Let us now suppose a second variate y and assume that the correlation of x and y for practical purposes is linear. Then clearly since a given x will have a constant class-mark, the correlation of y and C_x for a constant x is zero; that is to say that the partial correlation coefficient

$${}_x\rho_{y\,C_x}=\frac{r_{y\,C_x}-r_{xy}\,r_{x\,C_x}}{\sqrt{1-r^2_{xy}}\,\sqrt{1-r^2_{x\,C_x}}}=0.$$

Hence $$r_{yC_x} - r_{xy}\, r_{xC_x} = 0,$$

or $$r_{xy} = \frac{r_{yC_x}}{r_{xC_x}} \qquad \text{(iv)}.$$

In other words, to find the true correlation of x and any other variate y, divide the correlation of y and the class-mark of x, by the correlation of x with its class-mark.

From (iv) we can deduce the correction for number of arrays when we find η_{xy} the correlation ratio of x and y on the supposition that r_{xy} equals sufficiently closely η_{xy}, i.e. when we "find r by η methods." Let H_{yC_x} equal the value of η found when y is finely classified and the mean of an array can be determined for y, but the arrays of x are broad classes*. Then

$$\eta_{xy} = \frac{H_{yC_x}}{r_{xC_x}} = \frac{H_{yC_x}}{\sqrt{S(n_s \bar{x}_s^2)/(N\sigma_x^2)}} \qquad \text{(v)},$$

$$= \frac{H_{yC_x}}{\sqrt{S\dfrac{N}{n_s}(z_{s_1} - z_{s_2})^2}} \qquad \text{(vi)},$$

if we suppose the distribution normal.

(4) Now let us consider two variables x and y given by their class-marks C_x and C_y. If we correlate y with any variable u we have at once by (iv)

$$r_{uy} = \frac{r_{C_y u}}{r_{yC_y}} \qquad \text{(vii)}.$$

Now as u is quite arbitrary put it equal to the class-mark of x or C_x, then

$$r_{yC_x} = \frac{r_{C_y C_x}}{r_{yC_y}} \qquad \text{(viii)}.$$

Substitute this in (iv) and we have†

$$r_{xy} = \frac{r_{C_y C_x}}{r_{xC_x}\, r_{yC_y}} \qquad \text{(ix)}.$$

* The primary establishment of equation (v) is due to "Student." His paper published in this number reached me, as I was writing this paper. I had obtained (iv) and (ix) and used them to correct contingency, but not to correct η.

† A somewhat different proof of this formula may be obtained as follows: the partial correlation ${}_{xy}r_{C_x C_y}$ is clearly zero, for when x and y are constant C_x and C_y do not vary. Hence

$$r_{C_x C_y}(1 - r^2_{xy}) - r_{xC_x} r_{xC_y} - r_{yC_y} r_{yC_x} + r_{xy}(r_{xC_x} r_{yC_y} + r_{xC_y} r_{yC_x}) = 0.$$

But $$r_{xy} = r_{yC_x}/r_{xC_x} = r_{xC_y}/r_{yC_y},$$

and if we substitute for r_{yC_x} and r_{xC_y} we have

$$r_{C_x C_y}(1 - r^2_{xy}) - (r_{xy} - r^2_{xy})\, r_{yC_y} r_{xC_x} = 0,$$

or $$r_{xy} = \frac{r_{C_x C_y}}{r_{yC_y} r_{xC_x}}, \text{ as before.}$$

Thus $1/(r_{xC_x} r_{yC_y})$ appears to be the corrective factor when we group the variates x and y both into broad categories, the ranges of which are of any nature.

We have

$$r_{C_y C_x} = \frac{S(n_{st}\bar{x}_{st}\bar{y}_{st})/N}{\sqrt{S(n_s\bar{x}_s^2)/N}\sqrt{S(n_t\bar{y}_t^2)/N}} \quad \text{.....................(x)},$$

and

$$r_{xy} = \frac{S\left(\frac{n_{st}}{N}\frac{\bar{x}_{st}}{\sigma_x}\frac{\bar{y}_{st}}{\sigma_y}\right)}{S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) \times S\left(\frac{n_t}{N}\frac{\bar{y}_t^2}{\sigma_y^2}\right)} \quad \text{...........................(xi)},$$

where n_{st} is the frequency of the cell in the sth category of x and tth category of y.

It is obvious that r_{xy} takes its usual value when the classes of x and y become very numerous, for both factors in the denominator are then unity.

In the particular case of normal distributions, on the assumption that the product $\bar{x}_{st}\bar{y}_{st}$ is sufficiently close to $\bar{x}_s\bar{y}_t$ to be replaced by it:

$$r_{xy} = \frac{S\left\{\frac{n_{st}N}{n_s n_t}(z_{s_1} - z_{s_2})(z_{t_1} - z_{t_2})\right\}}{S\left\{\frac{N}{n_s}(z_{s_1} - z_{s_2})^2\right\} S\left\{\frac{N}{n_t}(z_{t_1} - z_{t_2})^2\right\}} \quad \text{..................(xii)},$$

which admits of fairly ready determination from Sheppard's Tables.

(5) An approximate value of r_{xC_x} can be reached in the following manner, if we suppose the range of the broad classes are all equal and given by h.

Assume a parabola $\qquad y = a + bx + cx^2$

to give by its area the three class frequencies n_{s-1}, n_s, n_{s+1}, the origin being at the start of the n_{s-1} class. Its equation will be found to be

$$y = \frac{2n_{s+1} - 7n_s + 11n_{s-1}}{6h} - \frac{n_{s+1} - 3n_s + 2n_{s-1}}{h}\left(\frac{x}{h}\right) + \frac{n_{s+1} - 2n_s + n_{s-1}}{2h}\left(\frac{x}{h}\right)^2 \quad \text{... (xiii)}.$$

But

$$n_s\bar{x}_s' = \int_h^{2h} yx\,dx = \frac{h}{24}(n_{s+1} + 36n_s - n_{s-1}),$$

or

$$\bar{x}_s' = \frac{3}{2}h + \frac{h}{24}\frac{n_{s+1} - n_{s-1}}{n_s}.$$

Thus, if x_s' be the distance from the origin to the *middle* of the sth class,

$$\bar{x}_s' - x_s' = \frac{1}{24}h\frac{n_{s+1} - n_{s-1}}{n_s},$$

and if $\bar{x}_s$ and x_s be measured from the mean of the whole population

$$\bar{x}_s = x_s - \frac{1}{24}h\frac{n_{s-1} - n_{s+1}}{n_s},$$

and

$$n_s\bar{x}_s^2 = n_s x_s^2 - \frac{n_s}{12} h x_s \frac{n_{s-1} - n_{s+1}}{n_s} + \frac{h^2}{576} \frac{(n_{s-1} - n_{s+1})^2}{n_s}$$

$$= n_s x_s^2 - \frac{h}{12}(x_{s-1}n_{s-1} - x_{s+1}n_{s+1}) - \frac{h^2}{12} n_{s-1} - \frac{h^2}{12} n_{s+1} + \frac{h^2}{576} \frac{(n_{s-1} - n_{s+1})^2}{n_s}.$$

Summing for all classes and dividing by $N\sigma_x^2$

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) = \frac{S(n_s x_s^2)}{N\sigma_x^2} - \frac{1}{12}\frac{h^2}{\sigma_x^2} - \frac{1}{12}\frac{h^2}{\sigma_x^2} + \frac{h^2}{576\sigma_x^2} S \frac{(n_{s-1} - n_{s+1})^2}{N n_s}.$$

But by Sheppard's Theorem, with contact at the tails of the order we are supposing:

$$\sigma_x^2 = \frac{S(n_s x_s^2)}{N} - \frac{1}{12}h^2.$$

Further, if we suppose y_{s-1}, y_s, y_{s+1}, y_{s+2} to be the bounding ordinates of the classes n_{s-1}, n_s, n_{s+1} of the frequency curve,

$$\alpha_{s-1} = n_{s-1} - y_s h = n_{s-1} - \frac{(-n_{s+1} + 5n_s + 2n_{s-1})}{6} = \frac{1}{6}(4n_{s-1} - 5n_s + n_{s+1}).$$

Similarly $\qquad \alpha_{s+1} = n_{s+1} - y_{s+2}h = \frac{1}{6}(n_{s-1} - 5n_s + 4n_{s+1}).$

Here α_{s-1} and α_{s+1} are the excesses of n_1 and n_3 above rectangles; and we have

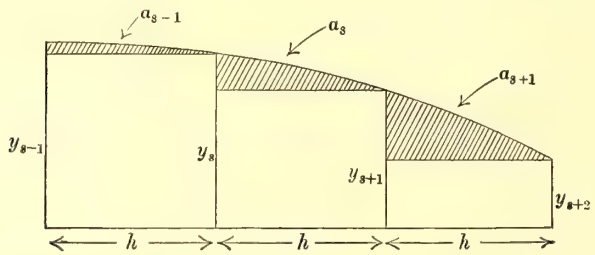
$$\alpha_{s-1} - \alpha_{s+1} = \tfrac{1}{2}(n_{s-1} - n_{s+1}).$$

Thus finally

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) = 1 - \frac{1}{12}\frac{h^2}{\sigma_x^2} + \frac{h^2}{144\sigma_x^2} S \frac{(\alpha_{s-1} - \alpha_{s+1})^2}{N n_s} \quad \text{............(xiv).}$$

Consider the last term; it may be written

$$S\left\{\frac{\frac{1}{2}(\alpha_s - \alpha_{s-1} + \alpha_{s+1} - \alpha_s)}{n_s} \times \frac{\frac{1}{2}(\alpha_s - \alpha_{s-1} + \alpha_{s+1} - \alpha_s)}{N}\right\} \times \frac{h^2}{36\sigma_x^2}.$$



Now the numerators are mean differences of what will be, if h be at all small, small areas, and these mean differences are divided in the one case by n_s and the other by N, much larger quantities. Hence the last term is as a rule very much smaller than the second, and we have approximately

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) = 1 - \frac{1}{12}\frac{h^2}{\sigma_x^2} \quad \text{..........................(xv).}$$

Thus we see how rapidly, if h be small with regard to σ_x, the value of the quantity

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right),$$

and therefore of r_{xC_x}, its square root, approaches unity.

(6) It is desirable to illustrate the approach of r_{xC_x} to unity as we increase the number of groups. I have therefore worked out this approach for three and more *symmetrical* groups when (a) we use the approximate formula (xv), (b) we still suppose the classes to have their ranges equal but the frequencies to be given by a normal curve, (c) we assume the frequencies in the classes and not the class ranges to be equal, the frequency being supposed normal, (d) we suppose the frequencies to increase by 50 % at each stage, i.e. to be as 1, 3/2, 9/4, 27/8, etc.

These cases will be sufficient to indicate what sort of frequencies we should take for few classes in order to get the highest correlation between variate and its class-mark. I shall after these theoretical investigations consider a few actual cases of "broad" categories.

Values of r_{xC_x} for Various Groupings.*

No. of Classes	(a) Equal Ranges, Any Frequency. Formula (xv)	(b) Equal Ranges, Normal Frequency	(c) Equal Subfrequencies, Normal Frequency	(d) Increasing Subfrequencies, Normal Frequency
3	·817 ($h=2\sigma$)	·859	·891	·876
4	·901 ($h=1·5\sigma$)	·915	·928	·912
5	·938 ($h=1·2\sigma$)	·943	·947	·928
6	·957 ($h=\sigma$)	·960	·959	·939
8	·978 ($h=·75\sigma$)	·977	·972	—
10	·985 ($h=·6\sigma$)	·985	·979	—
12	·990 ($h=·5\sigma$)	·989	·984	—
14	·992 ($h=\frac{3}{7}\sigma$)	·992	·987	—
16	·994 ($h=\frac{3}{8}\sigma$)	·994	·990	—
20	·996 ($h=·3\sigma$)	·996	·992	—

It is clear from the fourth column that nothing whatever is gained by exaggerating the frequencies of the extreme or tail sections. Further:

(a) Equal frequencies are better than equal ranges up to about six classes. After six classes it is better to take equal ranges.

(b) After six classes the approximate formula (xv) is amply sufficient in the case of equal ranges to obtain the value of r_{xC_x}.

Thus we have the general rule that up to six broad categories it is desirable to make the frequencies of those classes approximately equal, but beyond this greater

* The number of classes given in (a) and (b) is based upon the range of 6σ covering for such data as we usually deal with, i.e. 500 to 1000 cases, the total frequency.

exactness will be reached by equal ranges. The differences, however, between equal ranges and equal frequencies are after six classes so small—within the value of the usual probable error of random sampling—that either method will give practically quite good results.

(7) It is interesting to note the relation of the present method to the usual correction for grouping in the value of the correlation. In that method we take the groups at their mid-points, and we do not correct the product, but only the standard deviations of the two variates by the usual Sheppard's correction, $\frac{1}{12}$th of the sub-range squared being subtracted from the raw second moment coefficient. Now we have

$$r_{xy} = S\,(n_{st} x_s y_t)/\sigma_x \sigma_y,$$

and we have to replace x_s, y_t by $\bar{x}_s$ and $\bar{y}_t$.

But we have seen that approximately with equal ranges

$$\bar{x}_s = x_s - \frac{1}{24} h \frac{n_{s-1} - n_{s+1}}{n_s},$$

$$\bar{y}_t = y_t - \frac{1}{24} k \frac{n_{t-1} - n_{t+1}}{n_t},$$

and
$$S\left\{\frac{n_s}{N}\left(\frac{\bar{x}_s}{\sigma_x}\right)^2\right\} = 1 - \frac{1}{12}\frac{h^2}{\sigma_x^2}, \qquad S\left\{\frac{n_t}{N}\left(\frac{\bar{y}_t}{\sigma_y}\right)^2\right\} = 1 - \frac{1}{12}\frac{k^2}{\sigma_y^2}.$$

Now

$$S\,(n_{st}\,\bar{x}_s\,\bar{y}_t) = S\,(n_{st}\,x_s\,y_t) - \frac{1}{24} h S\left\{\frac{n_{s-1} - n_{s+1}}{n_s} n_{st} y_t\right\}$$
$$- \frac{1}{24} k S\left\{\frac{n_{t-1} - n_{t+1}}{n_t} n_{st} x_s\right\} + \frac{1}{576} hk S\left\{n_{st} \frac{n_{s-1} - n_{s+1}}{n_s} \frac{n_{t-1} - n_{t+1}}{n_t}\right\} \quad \text{......(xvi)}.$$

Consider first the last term*, it contains not only the product of hk, but also the product of differences, and is of the form, when we divide by $N\sigma_x\sigma_y$,

$$\frac{1}{36}\frac{h}{\sigma_x}\frac{k}{\sigma_y} S\left\{\frac{n_{st}}{N} \times \frac{\frac{1}{2}(\alpha_s - \alpha_{s-1}) + \frac{1}{2}(\alpha_{s+1} - \alpha_s)}{n_s} \times \frac{\frac{1}{2}(\alpha'_t - \alpha'_{t-1}) + \frac{1}{2}(\alpha'_{t+1} - \alpha'_t)}{n_t}\right\},$$

which will be of the fourth order of small quantities if h/σ_x and k/σ_y be small and may be neglected as compared to the square of small quantities.

* As a matter of fact for Gaussian frequency

$$S\left(n_{st}\frac{n_{s-1} - n_{s+1}}{n_s}\frac{n_{t-1} - n_{t+1}}{n_t}\right) = 4hk\,S\,(n_{st}\bar{x}_s\bar{y}_t)$$

to our degree of approximation. Thus the fourth term may be written $\frac{1}{144} h^2k^2\,S\,(n_{st}x_sy_t)$, which gives

$$S\,(n_{st}\bar{x}_s\bar{y}_t) = S\,(n_{st}x_sy_t)\left(1 - \frac{1}{12}h^2 - \frac{1}{12}k^2 + \frac{1}{144}h^2k^2\right)$$
$$= S\,(n_{st}x_sy_t)\left(1 - \frac{1}{12}h^2\right)\left(1 - \frac{1}{12}k^2\right).$$

Next, considering the second summation and summing first $n_{st}y_t$ for all values of t, it equals $n_s\bar{y}_s$, where $\bar{y}_s$ is mean of all y's in array corresponding to x_s. But if the regression be linear $\bar{y}_s = \frac{r\sigma_y}{\sigma_x} x_s$. Thus the second sum equals

$$\frac{1}{24} r\sigma_y\sigma_x \frac{h}{\sigma_x^2} S\{(n_{s-1} - n_{s+1}) x_s\}$$

$$= \frac{1}{24} r\sigma_y\sigma_x \frac{h}{\sigma_x^2} \{S(x_{s-1}n_{s-1} - x_{s+1}n_{s+1}) + hS(n_{s-1}) + hS(n_{s+1})\}$$

$$= \frac{1}{24} r\sigma_y\sigma_x \frac{h^2}{\sigma_x^2} 2N$$

as before. But in a term of this order we may put

$$Nr\sigma_y\sigma_x = S(n_{st}x_s y_t).$$

Similarly the third summation

$$= \frac{1}{12} S(n_{st}x_s y_t) \times \frac{k^2}{\sigma_y^2}.$$

Or, finally,

$$S(n_{st}\bar{x}_s\bar{y}_t) = S(n_{st}x_s y_t)\left(1 - \frac{h^2}{12\sigma_x^2} - \frac{k^2}{12\sigma_y^2}\right) \quad \text{.........(xvi) } bis.$$

Thus

$$\frac{S\left(\frac{n_{st}}{N}\frac{\bar{x}_s}{\sigma_x}\frac{\bar{y}_t}{\sigma_y}\right)}{S\left(\frac{n_s\bar{x}_s^2}{N\sigma_x^2}\right)S\left(\frac{n_t\bar{y}_t^2}{N\sigma_y^2}\right)} = \frac{\frac{1}{\sigma_x\sigma_y}S\left(\frac{n_{st}}{N}x_s y_t\right)\left(1 - \frac{h^2}{12\sigma_x^2} - \frac{k^2}{12\sigma_y^2}\right)}{\left(1 - \frac{1}{12}\frac{h^2}{\sigma_x^2}\right)\left(1 - \frac{1}{12}\frac{k^2}{\sigma_y^2}\right)},$$

or, since we neglect terms of fourth order,

$$\frac{S\left(\frac{n_{st}}{N}\frac{\bar{x}_s}{\sigma_x}\frac{\bar{y}_t}{\sigma_y}\right)}{S\left(\frac{n_s\bar{x}_s^2}{N\sigma_x^2}\right)S\left(\frac{n_t\bar{y}_t^2}{N\sigma_y^2}\right)} = \frac{S\left(\frac{n_{st}}{N}x_s y_t\right)}{\sigma_x\sigma_y} = r_{xy} \quad \text{...............(xvii)},$$

the usual value, σ_x and σ_y having of course to be corrected by Sheppard.

We see by (xv) that

$$\sigma_x^2 = S\left(\frac{n_s}{N}\bar{x}_s^2\right)\left(1 + \frac{1}{12}\frac{h^2}{\sigma_x^2}\right) \quad \text{..................(xvii) } bis,$$

but by Sheppard

$$\sigma_x^2 = S\left(\frac{n_s}{N}x_s^2\right)\left(1 - \frac{1}{12}\frac{h^2}{\sigma_x^2}\right).$$

Accordingly

$$S\left(\frac{n_s}{N}\bar{x}_s^2\right) = S\left(\frac{n_s}{N}x_s^2\right)\left(1 - \frac{1}{6}\frac{h^2}{\sigma_x^2}\right) \quad \text{..................(xviii)},$$

which enables us with equal small subranges to use the standard deviation of means or mid-abscissae at our pleasure.

(8) Of course it is absurd in practice to push our results to the extreme* of two categories only, but theoretically it is not without interest to note the results which flow from such an assumption.

* We have seen that $\bar{x}_{st}\bar{y}_{st}$ can only be replaced by $\bar{x}_s\bar{y}_t$ in (xi), or as a special case, (xii) used, provided the subranges are equal and fairly small.

Let the division be into two categories containing $N \times \frac{1}{2}(1+\alpha)$ and $N \times \frac{1}{2}(1-\alpha)$ individuals, and let no assumption be made as to the nature of the frequency. Then

$$S\left(\frac{n_s}{N}\bar{x}_s^2\right) = \tfrac{1}{2}(1+\alpha)\bar{x}_1^2 + \tfrac{1}{2}(1-\alpha)\bar{x}_2^2,$$

but
$$\tfrac{1}{2}(1+\alpha)\bar{x}_1 + \tfrac{1}{2}(1-\alpha)\bar{x}_2 = 0,$$

and therefore
$$S\left(\frac{n_s}{N}\bar{x}_s^2\right) = \bar{x}_2^2\,\frac{\frac{1}{2}(1-\alpha)}{\frac{1}{2}(1+\alpha)}.$$

Thus
$$r_{xC_x} = \frac{\bar{x}_2}{\sigma_x}\sqrt{\frac{\frac{1}{2}(1-\alpha)}{\frac{1}{2}(1+\alpha)}} \quad \text{..........................(xix).}$$

We note therefore that unless we suppose all the frequency in each category concentrated at one point—a very rare occurrence—r_{xC_x} will always be less than unity, and therefore in correlating class indices a correction will always have to be made. We will consider two cases: (i) when the frequency distribution is a rectangle of length l. In this case $\bar{x}_2$ is $= l\{\frac{1}{2}(1+\alpha)\}$ and $\sigma_x = l/\sqrt{3}$, we have then

$$r_{xC_x} = \sqrt{3 \times \tfrac{1}{2}(1+\alpha) \times \tfrac{1}{2}(1-\alpha)} \quad \text{.....................(xx);}$$

(ii) the distribution of frequency is supposed to be Gaussian. In this case

$$\bar{x}_2 = \frac{z_\alpha \times \sigma_x}{\frac{1}{2}(1-\alpha)},$$

and
$$r_{xC_x} = \frac{z_\alpha}{\sqrt{\frac{1}{2}(1-\alpha) \times \frac{1}{2}(1+\alpha)}} \quad \text{........................(xxi),}$$

in other words, r_{xC_x} is the reciprocal of the χ_α already tabled in this number of *Biometrika*, p. 27.

We have the following results:

Value of $\frac{1}{2}(1+\alpha)$ as percentage of total	Value of r_{xC_x}	
	Gaussian Frequency	Rectangular Frequency
50 %	·798	·866
60 %	·789	·849
70 %	·759	·794
80 %	·700	·693
90 %	·585	·520
95 %	·473	·377
99 %	·268	·172

It will be seen at once how rapidly the corrective factor $1/r_{xC_x}$ rises as the division of the categories becomes more and more unequal. But in both cases very sensible correction is needful and its nature depends on the particular frequency.

Let us now pass to the consideration of the correlation r_{xy} as obtained by correction from the class index correlation, i.e.

$$r_{xy} = \frac{r_{C_x C_y}}{r_{x C_x} \times r_{y C_y}}.$$

The difficulty is to know how to determine the product $S(n_{st}\bar{x}_{st}\bar{y}_{st})$ for such a slender division as a fourfold table. If we assume it still to be equal to

$$S(n_{st}\bar{x}_s\bar{y}_t)$$

we can then find r_{xy}.

We have $$S(n_{st}\bar{x}_s\bar{y}_t) = n_1\bar{x}_1\bar{y}_1 + n_2\bar{x}_2\bar{y}_2 + n_3\bar{x}_3\bar{y}_3 + n_4\bar{x}_4\bar{y}_4,$$

where we take for our fourfold table:

n_1	n_2	n_1+n_2
n_4	n_3	n_3+n_4
n_1+n_4	n_2+n_3	N

But clearly on the above assumption

$$\bar{x}_1 = \bar{x}_4, \quad \bar{x}_2 = \bar{x}_3, \quad \bar{y}_1 = \bar{y}_2, \quad \bar{y}_3 = \bar{y}_4.$$

Thus $$S(n_{st}\bar{x}_s\bar{y}_t) = n_1\bar{x}_1\bar{y}_1 + n_2\bar{x}_2\bar{y}_1 + n_3\bar{x}_2\bar{y}_4 + n_4\bar{x}_1\bar{y}_4,$$

then since $$(n_1+n_4)\bar{x}_1 + (n_2+n_3)\bar{x}_2 = 0,$$

$$(n_1+n_2)\bar{y}_1 + (n_3+n_4)\bar{y}_4 = 0,$$

we easily deduce

$$S(n_{st}\bar{x}_s\bar{y}_t) = \frac{\bar{x}_2\bar{y}_4(n_1n_3 - n_3n_1)}{(n_1+n_4)(n_1+n_2)} \quad \text{.....................(xxii).}$$

Similarly we find

$$\left.\begin{aligned} \sigma^2_{\bar{x}_s} &= \frac{n_2+n_3}{n_1+n_4}\bar{x}_2^2 \\ \sigma^2_{\bar{y}_t} &= \frac{n_3+n_4}{n_1+n_2}\bar{y}_4^2 \end{aligned}\right\} \quad \text{...........................(xxiii).}$$

And accordingly

$$r_{C_x C_y} = \frac{S(n_{st}\bar{x}_s\bar{y}_t)}{\sigma_{\bar{x}}\sigma_{\bar{y}}}$$

$$= \frac{n_1n_3 - n_2n_4}{\sqrt{(n_1+n_2)(n_1+n_4)(n_2+n_3)(n_3+n_4)}} \quad \text{............(xxiv).}$$

This is not really the true value of the class index correlation unless we assume $\bar{x}_s\bar{y}_t = \bar{x}_{st}\bar{y}_{st}$, which is not generally true. It is really the correlation of the means used as class-marks, and as shown long ago by me it is the correlation of errors in the means of the two variates, and also equals ϕ the basis of the mean square contingency. If we correct to obtain r_{xy} by the factors

$$r_{x C_x} = \sigma_{\bar{x}_s}/\sigma_x \quad \text{and} \quad r_{y C_y} = \sigma_{\bar{y}_t}/\sigma_y,$$

we find

$$r_{xy} = \frac{n_1 n_3 - n_2 n_4}{(n_2 + n_3)(n_3 + n_4) \times \dfrac{\bar{x}_2}{\sigma_x} \times \dfrac{\bar{y}_4}{\sigma_y}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxv)}.$$

Mr Yule has termed (xxiv) the "theoretical value of the correlation*." It will be quite obvious from the present discussion that it is only an approximate value even to the class index correlation, and cannot represent r_{xy} at all, unless the frequency distributions each consist of two isolated points. Generally

$$r_{xy} = \frac{n_1 n_3 - n_3 n_1}{\sqrt{(n_1 + n_2)(n_1 + n_4)(n_2 + n_3)(n_3 + n_4)}} \times \sqrt{\frac{\sigma_x}{\bar{x}_1} \frac{\sigma_x}{\bar{x}_2} \frac{\sigma_y}{\bar{y}_1} \frac{\sigma_y}{\bar{y}_4}}$$

$$= r_{C_x C_y} \times \sqrt{\frac{\sigma_x}{\bar{x}_1} \cdot \frac{\sigma_x}{\bar{x}_2} \cdot \frac{\sigma_y}{\bar{y}_1} \cdot \frac{\sigma_y}{\bar{y}_4}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxvi)}.$$

This corrective value on $r_{C_x C_y}$ is always true, although the above value of $r_{C_x C_y}$ is only approximate.

This is brought out at once by considering what (xxv) reduces to in the case of Gaussian frequency. Here

$$(n_2 + n_3)\,\bar{x}_2 = N z_2 \sigma_x = N \sigma_x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{h^2}{\sigma_x^2}},$$

$$(n_3 + n_4)\,\bar{y}_4 = N z_4 \sigma_y = N \sigma_y \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{k^2}{\sigma_y^2}},$$

where h and k are the distances of the dividing lines from the mean. Hence if H and K have their usual meanings (xxv) gives us

$$r_{xy} = \frac{n_1 n_3 - n_2 n_4}{N^2 HK} = \epsilon \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xxvii)}.$$

ϵ is the expression used in my memoir on the correlation of characters not quantitatively measurable†. Thus the physical meaning of ϵ is $r_{C_x C_y}$ corrected for the use of class indices, but not for the assumption that $\bar{x}_{st} \times \bar{y}_{st}$ may be replaced by $\bar{x}_s \times \bar{y}_t$. We may therefore anticipate that ϵ or, failing Gaussian frequency, (xxv), will give a better approximation than $r_{C_x C_y}$ to the true value of the correlation of x and y.

I shall discuss elsewhere the possibility of any absolute identity in the members of a class; the only thing in which I have personally come across it is in theoretical investigations—never in practical—on Mendelian units. When we are absolutely ignorant of the nature of the frequency, then I feel sure that in dealing with either physical or mental characteristics in living forms, the Gaussian distribution will in the long run be found closer than any single type of distribution hitherto discussed, and that accordingly even ϵ—to say nothing

* *Introduction to the Theory of Statistics*, p. 212.

† *Phil. Trans.* Vol. 195, A, p. 7.

of r by fourfold—would be far safer to use than $r_{C_xC_y}$, or any arbitrary coefficient of association, which wholly neglects the question of the type of the frequency distribution. In the rare cases in which the frequency is collected into points—absolute homogeneity of category contents—then $r_{C_xC_y}$ is the right coefficient to use and not Yule's coefficient of association.

(9) I propose in this section to consider the relations between

$$S\left(\frac{n_{st}}{N}\cdot\bar{x}_{st}\cdot\bar{y}_{st}\right), \quad S\left(\frac{n_{st}}{N}\cdot\bar{x}_s\cdot\bar{y}_s\right) \text{ and } S\left(\frac{n_{st}}{N}\cdot x_s\cdot y_s\right)$$

when the subranges for each character are small and equal. Let those for x be h, and for y be k. Then consider the surface

$$z = a + bx + cy + dx^2 + ey^2,$$

and choose a, b, c, d, e so that it gives by its volumes correctly the five class frequencies

$$n_{s-1,\,t}, \quad n_{s,\,t}, \quad n_{s+1,\,t}, \quad n_{s,\,t-1}, \quad n_{s,\,t+1}.$$

Then if we take the origin so that it lies at the mid-point of the $n_{s,\,t}$ group, we find

$$\begin{aligned} z = \frac{1}{hk}\left\{n_{s,\,t} - \frac{1}{12}(n_{s-1,\,t} + n_{s,\,t-1} - 2n_{s,\,t}) - \frac{1}{24}(n_{s+1,\,t} - n_{s-1,\,t} + n_{s,\,t+1} - n_{s,\,t-1})\right\} \\ + \frac{1}{2}\frac{x}{h}\frac{n_{s+1,\,t} - n_{s-1,\,t}}{hk} + \frac{1}{2}\frac{y}{k}\frac{n_{s,\,t+1} - n_{s,\,t-1}}{hk} \\ + \frac{x^2}{h^2}\left(\frac{n_{s-1,\,t} - n_{s,\,t}}{hk} + \frac{n_{s+1,\,t} - n_{s,\,t}}{2hk}\right) + \frac{y^2}{k^2}\left(\frac{n_{s,\,t-1} - n_{s,\,t}}{hk} + \frac{n_{s,\,t+1} - n_{s,\,t}}{2hk}\right). \end{aligned}$$

Further

$$n_{st}(\bar{x}_{st} - x_s) = \int_{-\frac{1}{2}k}^{+\frac{1}{2}k}\int_{-\frac{1}{2}h}^{+\frac{1}{2}h} xz\,dx = \frac{1}{12}bh^3k = \frac{h^3k}{24}\left(\frac{n_{s+1,\,t} - n_{s-1,\,t}}{h^2k}\right).$$

Or

$$\bar{x}_{st} = x_s + \frac{1}{24}h\frac{n_{s+1,\,t} - n_{s-1,\,t}}{n_{st}} \quad \text{.........................(xxviii).}$$

Similarly

$$\bar{y}_{st} = y_t + \frac{1}{24}k\frac{n_{s,\,t+1} - n_{s,\,t-1}}{n_{st}} \quad \text{.........................(xxix).}$$

Now take the product of these, multiply by n_{st} and sum, we have

$$\begin{aligned} S\left(\frac{n_{st}}{N}\bar{x}_{st}\bar{y}_{st}\right) = S\left(\frac{n_{st}}{N}x_s y_t\right) + \frac{1}{24}hS\left\{\frac{y_t}{N}(n_{s+1,\,t} - n_{s-1,\,t})\right\} \\ + \frac{1}{24}kS\left\{\frac{x_s}{N}(n_{s,\,t+1} - n_{s,\,t-1})\right\} \\ + \frac{1}{144}hkS\left(\frac{n_{s,\,t+1} - n_{s,\,t} + n_{s,\,t} - n_{s,\,t-1}}{2N}\right) \times \left(\frac{n_{s+1,\,t} - n_{s,\,t} + n_{s,\,t} - n_{s-1,\,t}}{2n_{st}}\right). \end{aligned}$$

But if h and k be small the last summation is, precisely as that dealt with on p. 122, of the fourth order, and we have

$$S\left(\frac{n_{st}}{N}\bar{x}_{st}\bar{y}_{st}\right) = S\left(\frac{n_{st}}{N}x_s y_t\right) \quad \text{.........................(xxx),}$$

for the second and third sums obviously vanish if we suppose zero frequencies at the boundaries of each row or column. Hence by (xvi) *bis* we have

$$S\left(\frac{n_{st}}{N}\bar{x}_{st}\bar{y}_{st}\right) = S(n_{st}\bar{x}_s\bar{y}_t) \times \left(1 + \frac{h^2}{12\sigma_x^2} + \frac{k^2}{12\sigma_y^2}\right) \quad \text{.........(xxxi).}$$

Thus we cannot replace $S(n_{st}\bar{x}_{st}\bar{y}_{st})$ by $S(n_{st}\bar{x}_s\bar{y}_t)$ unless h and k are very small relative to σ_x and σ_y respectively.

But now consider $r_{xy} = \dfrac{S(n_{xy}xy)}{N\sigma_x\sigma_y}$, the accurate value.

We may replace $S(n_{xy}xy)$ by $S(n_{st}x_sy_s)$, there being no Sheppard's correction for the product moment. Hence

$$r_{xy} = \frac{S\left(\frac{n_{st}}{N}x_sy_s\right)}{\sigma_x\sigma_y} = \frac{S\left(\frac{n_{st}}{N}\bar{x}_{st}\bar{y}_{st}\right)}{\sigma_x\sigma_y} \text{ by (xxx)}$$

$$= S\left(\frac{n_{st}}{N}\frac{\bar{x}_s}{\sigma_x}\frac{\bar{y}_t}{\sigma_y}\right)\left(1 + \frac{h^2}{12\sigma_x^2} + \frac{k^2}{12\sigma_y^2}\right) \text{ by (xxxi)}$$

$$= \frac{S\left(\frac{n_{st}}{N}\frac{\bar{x}_s}{\sigma_x}\frac{\bar{y}_t}{\sigma_y}\right)}{\left(1 - \frac{h^2}{12\sigma_x^2}\right)\left(1 - \frac{k^2}{12\sigma_y^2}\right)} \text{ to the same degree of approximation}$$

$$= \frac{S\left(\frac{n_{st}}{N}\frac{\bar{x}_s}{\sigma_x}\frac{\bar{y}_t}{\sigma_y}\right)}{S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) \times S\left(\frac{n_t}{N}\frac{\bar{y}_t^2}{\sigma_y^2}\right)} \text{ by (xvii) } bis \quad \text{..............................(xxxii).}$$

This formula is free from all Sheppard's corrections, and *in form* it does not involve the equality of the subranges. Hence it seems likely to give moderately good results even for unequal subranges, provided they do not differ very widely, and there are not too few of them. It leads as before in the case of normal frequency to

$$r_{xy} = \frac{S\left\{\frac{Nn_{st}}{n_sn_t}(z_{s_1} - z_{s_2})(z_{t_1} - z_{t_2})\right\}}{S\left\{\frac{N}{n_s}(z_{s_1} - z_{s_2})^2\right\} S\left\{\frac{N}{n_t}(z_{t_1} - z_{t_2})^2\right\}},$$

but it is clear that this cannot be pressed so far as to make only a very few, and those very unequal, divisions for each variate.

(10) The present method of correction seems likely to give good results in the case of the method of mean square contingency*, when we may reasonably assume the distribution not widely divergent from the Gaussian.

* A much fuller discussion of all the corrections for contingency has been some years in progress and will shortly appear.

Illustration I. Thus I take the table for correlation of stature in father and son, the subranges being each two inches.

Stature of Father.

Stature of Son.	58·5—60·5	60·5—62·5	62·5—64·5	64·5—66·5	66·5—68·5	68·5—70·5	70·5—72·5	72·5—74·5	74·5—76·5	Totals
59·5—61·5	—	—	1·5	1	1	—	—	—	—	3·5
61·5—63·5	0·5	2·75	5·75	9·5	5	0·25	0·25	—	—	24
63·5—65·5	4	7·75	20	41·5	17·25	8·25	1·25	—	—	100
65·5—67·5	2	10	32	73	78·75	33·5	7·25	1	—	237·5
67·5—69·5	—	4·5	27·75	65·5	95	93·25	31·5	4·5	1	323
69·5—71·5	—	—	6·75	38·25	61	77·5	39·5	11	2	236
71·5—73·5	—	—	0·25	5·75	24·75	34·5	32·25	7	0·5	105
73·5—75·5	—	—	1	3	6·25	6·75	13	5·5	2	37·5
75·5—77·5	—	—	—	—	2·5	1·5	1·5	2·5	—	8
77·5—79·5	—	—	—	—	—	2	0·5	1	—	3·5
Totals	6·5	25	95	237·5	291·5	257·5	127	32·5	5·5	1078

Worked out by the product-moment method and using the Sheppard correction for the S.D.'s, $r_{xy} = \cdot 5259$.

I then proceed to make it into two further tables (i) a 5×5, and (ii) a 3×3 celled table. These are as follows:

Stature of Father.

Stature of Son.	58·5—64·5	64·5—66·5	66·5—68·5	68·5—70·5	70·5—76·5	Totals
59·5—65·5	42·25	52	23·25	8·5	1·5	127·5
65·5—67·5	44	73	78·75	33·5	8·25	237·5
67·5—69·5	32·25	65·5	95	93·25	37	323
69·5—71·5	6·75	38·25	61	77·5	52·5	236
71·5—79·5	1·25	8·75	33·5	44·75	65·75	154
Totals	126·5	237·5	291·5	257·5	165	1078

and

Stature of Father.

Stature of Son.	*58·5—66·5*	*66·5—68·5*	*68·5—76·5*	Totals
59·5—67·5	211·25	102	51·75	365
67·5—69·5	97·75	95	130·25	323
69·5—79·5	55	94·5	240·5	390
Totals	364	291·5	422·5	1078

In obtaining these tables I was generally guided by an endeavour to roughly equalise the frequencies in the totals.

If x be the stature of father and y of son, then for the 5 × 5-fold table we have, reading from left to right and from top to bottom:

$$\bar{x}_s = -1{\cdot}6780,\quad -{\cdot}7650,\quad -{\cdot}0694,\quad +{\cdot}6191,\quad +1{\cdot}5440,$$
$$\bar{y}_t = -1{\cdot}6741,\quad -{\cdot}7617,\quad -{\cdot}0298,\quad +{\cdot}6811,\quad +1{\cdot}5795.$$

These give

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) = {\cdot}917{,}091, \qquad S\left(\frac{n_t}{N}\frac{\bar{y}_t^2}{\sigma_y^2}\right) = {\cdot}863{,}6599,$$
$$r_{x\,C_x} = {\cdot}9576, \qquad r_{y\,C_y} = {\cdot}9293.$$

I now calculated ϕ^2 and found for raw $N\phi^2$ the value 314·7514; deducting the correction for 5 × 5 cells, this equals 298·7514, whence $\phi^2 = {\cdot}277{,}1349$,

$$C = \sqrt{\frac{\phi^2}{1+\phi^2}} = {\cdot}4659.$$

Taking this to represent the correlation of C_x and C_y we should have

$$r_{xy} = {\cdot}4659/({\cdot}9576 \times {\cdot}9293)$$
$$= {\cdot}5235,$$

which is excellently in keeping with the value found from the 9 × 10 table by product moment, and not very different from the value ·5140 found from the original 17 × 20 table in inches by the same process.

I now turn to the 3 × 3-fold table and find the raw $N\phi^2 = 219{\cdot}2194$ or less the correction for 3 × 3 cells = 215·2194, whence $\phi^2 = {\cdot}199{,}6469$. Thus $C = {\cdot}40795$.

For this table reading in same directions as before:

$$\bar{x}_s = -1{\cdot}0823,\quad -{\cdot}0694,\quad +\ {\cdot}9803,$$
$$\bar{y}_t = -1{\cdot}0804,\quad -{\cdot}0298,\quad +1{\cdot}0358,$$

whence

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) = {\cdot}773{,}470, \qquad S\left(\frac{n_t}{N}\frac{\bar{y}_t^2}{\sigma_y^2}\right) = {\cdot}783{,}6384,$$
$$r_{x\,C_x} = {\cdot}8795, \qquad r_{y\,C_y} = {\cdot}8852.$$

Thus taking as before $$r_{xy} = C/(r_{x\,C_x} r_{y\,C_y}),$$

we find $$r_{xy} = {\cdot}5240.$$

Again in excellent agreement with the 5 × 5 contingency result and also the product-moment result. The evaluations of $\bar{x}_s$ and $\bar{y}_t$ are of course based on the assumption of a Gaussian distribution for those variates.

Illustration II. The following table is given by Mr W. H. Gilby in a paper in *Biometrika* (Vol. VIII., p. 106) on the "Teacher's Appreciation of General Intelligence." It correlates class of intelligence with character of the clothing in boys. By treating the grades of intelligence and clothing as Gaussian variates an almost linear regression line was obtained in that paper. Hence it seems a favourable case to examine the present corrections upon. Intelligence was

measured on my scale (see *Biometrika*, Vol. VIII., p. 93) and clothing was divided into five categories, IV and V being ultimately classed together. The better intelligence has the higher letter, the worse clothing the higher number. The table runs:

Intelligence.

Clothing.	*B*	*C*	*D*	*E*	*F*	*G*	Totals
I	33	48	113	209	194	39	636
II	41	100	202	255	138	15	751
III	39	58	70	61	33	4	265
IV and V	17	13	22	10	10	1	73
Totals	130	219	407	535	375	59	1725

I shall use x for Intelligence and y for Clothing. Mr Gilby found raw $\phi^2 = \cdot 1014$, corrected for 4×6 cells $\phi^2 = \cdot 0927$ and $C = \cdot 291$. I find from left to right:

$$\bar{x}_s = -1\cdot 8859, \quad -1\cdot 1009, \quad -\cdot 4758, \quad +\cdot 2428, \quad +1\cdot 1179, \quad +2\cdot 2167,$$

$$\bar{y}_t = -1\cdot 0066, \quad +\cdot 2173, \quad +1\cdot 2130, \quad +2\cdot 1317.$$

Hence

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) = \cdot 933{,}393, \qquad S\left(\frac{n_t}{N}\frac{\bar{y}_t^2}{\sigma_y^2}\right) = \cdot 812{,}475,$$

$$r_{x\,C_x} = \cdot 9661, \qquad r_{y\,C_y} = \cdot 9014.$$

If we now correct Mr Gilby's contingency for the class-index correlations of x and y we have

$$r_{xy} = C/(r_{x\,C_x} r_{y\,C_y}) = \cdot 3342.$$

This is very near the values, ·343 and ·340, found by converting the table into a three-rowed table, classes III, IV and V of clothing being grouped together and a bi-serial η method used (*Biometrika*, Vol. VIII. p. 98).

I then proceeded to work out on the full table given above the value of r as determined by the approximate formula (xvii). This can be conveniently arranged thus:

	$-1\cdot 8859$	$-1\cdot 1009$	$-\cdot 4758$	$+\cdot 2428$	$+1\cdot 1179$	$+2\cdot 2167$
$-1\cdot 0066$	33	48	113	209	194	39
$+\cdot 2173$	41	100	202	255	138	15
$+1\cdot 2130$	39	58	70	61	33	4
$+2\cdot 1317$	17	13	22	10	10	1
	$+33\cdot 2178$	$+48\cdot 3168$	$+113\cdot 7458$	$-210\cdot 3794$	$-195\cdot 2804$	$-39\cdot 2574$
	$-8\cdot 9093$	$-21\cdot 7300$	$-43\cdot 8946$	$+55\cdot 4115$	$+29\cdot 9874$	$+3\cdot 2595$
	$-47\cdot 3070$	$-70\cdot 3540$	$-84\cdot 9100$	$+73\cdot 9930$	$+40\cdot 0290$	$+4\cdot 8520$
	$-36\cdot 2389$	$-27\cdot 7121$	$-46\cdot 8974$	$+21\cdot 3170$	$+21\cdot 3170$	$+2\cdot 1317$
	$-59\cdot 2374$	$-71\cdot 4793$	$-61\cdot 9562$	$-59\cdot 6579$	$-103\cdot 9470$	$-29\cdot 0142$

Therefore $$S\left(\frac{n_{st}\bar{x}_s\bar{y}_t}{N\sigma_x\sigma_y}\right) = -\cdot 240{,}51548,$$

$$r_{C_x C_y} = -\frac{\cdot 240{,}51548}{\cdot 9661 \times \cdot 9014} = -\cdot 2762,$$

and $$r_{C_x C_y}/(r_{x C_x} r_{y C_y}) = -\cdot 3171^*.$$

In the above table each row is first multiplied by the y_t on the left and the *final* sign at once given to the product; this forms the lower half of the table. The columns are then added up with the results given at the foot. The column sums are then multiplied by their respective $\bar{x}_s$'s, shown at the top of the first half of the table, regardless of the sign of $\bar{x}_s$, and the sum divided by 1725 as a continuous operation on the calculator, which shows finally ·240,51548.

The value ultimately obtained, ·32, is somewhat less than the ·33 of the contingency result but of the same order for all practical purposes.

I next reduce my table to a 4 × 4 table as follows:

Intelligence.

Clothing.	*B*+*C*	*D*	*E*	*F*+*G*	Totals
I	81	113	209	233	636
II	141	202	255	153	751
III	97	70	61	37	265
IV and V	30	22	10	11	73
Totals	349	407	535	434	1725

We have $\bar{x}_s = -1\cdot 3934,\quad -\cdot 4758,\quad +\cdot 2428,\quad +1\cdot 2674,$

$\bar{y}_t = -1\cdot 0066,\quad +\cdot 2173,\quad +1\cdot 2130,\quad +2\cdot 1317,$

$$S\left(\frac{n_s\bar{x}_s^2}{N\sigma_x^2}\right) = \cdot 868{,}606, \qquad S\left(\frac{n_t\bar{y}_t^2}{N\sigma_y^2}\right) = \cdot 812{,}475,$$

$$r_{x C_x} = \cdot 9320, \qquad r_{y C_y} = \cdot 9014.$$

Proceeding as before by formula (xvii) we find

$$r_{xy} = \frac{-\cdot 228{,}76485}{\cdot 868{,}606 \times \cdot 812{,}475} = -\cdot 3242,$$

a result which is again in excellent agreement with the previous ones.

Lastly I convert my table into a 3 × 3 table:

Intelligence.

Clothing.	*B*+*C*+*D*	*E*	*F*+*G*	Totals
I	194	209	233	636
II	343	255	153	751
III+IV+V	219	71	48	338
Totals	756	535	434	1725

* The negative sign shows that worse intelligence is associated with the poorer clothing.

I found $\bar{x}_s = -\ \cdot 8995, \quad + \cdot 2428, \quad + 1\cdot 2674,$

$$\bar{y}_t = -1\cdot 0066, \quad + \cdot 2173, \quad + 1\cdot 4110,$$

and $$S\left(\frac{n_s \bar{x}_s^2}{N\sigma_x^2}\right) = \cdot 776{,}941, \qquad S\left(\frac{n_t \bar{y}_t^2}{N\sigma_y^2}\right) = \cdot 784{,}241,$$

$$r_{xC_x} = \cdot 8814, \qquad r_{yC_y} = \cdot 8856.$$

Proceeding exactly as before by formula (xvii) we have

$$r_{xy} = \frac{-\cdot 204{,}01275}{\cdot 784{,}241 \times \cdot 776{,}941} = -\cdot 3348.$$

For comparison I worked out the above 3×3 table by contingency. The contingency corrected for 3×3 cells is $\phi^2 = \cdot 076{,}0779$, and therefore $C = \cdot 2659$, hence

$$r_{xy} = \cdot 2659/(\cdot 8814 \times \cdot 8856)$$

$$= \cdot 3406, \text{ regardless of sign.}$$

To sum up our results for this case we have:

	Bi-Serial η	Corrected Contingency	Formula (xvii)
6×2 Table	·340 and ·343	—	—
6×4 Table	—	·3342	·3171
4×4 Table	—	—	·3242
3×3 Table	—	·3406	·3348

The results are very satisfactory and show that for correlations of this magnitude our corrective factors work excellently when we use contingency to find $r_{C_x C_y}$. Formula (xvii) also gives fairly consistent results, but this is partially due to the relative smallness of r. For the smaller r the more nearly $\bar{x}_{st} \times \bar{y}_{st}$ may be replaced by $\bar{x}_s \times \bar{y}_t$ as is well shown in the case of the simple 2×2 table.

On my scale of *Intelligence* seven categories are used; there were in Mr Gilby's schools none of the A or Mentally Defective class, so that he only used six, B—G. We see that there is a high degree of correlation between the actual intelligence of a child and its class index. We must have

For seven classes		$r > \cdot 97$,
„ six	„	$r = \cdot 97$,
„ four	„	$r = \cdot 93$,
„ three	„	$r = \cdot 88$.

It is of interest to put beside these the frequencies of my General Health Scale and the resulting class index correlations. The scale had six classes:

Very Robust, Robust, Normally Healthy, Rather Delicate, Delicate and Very Delicate. Applied to 2037 schoolgirls (*Biometrika,* Vol. III. p. 166) it gave:

	V. R.	R.	N. H.	R. D. + D.	V. D.	Total
	109	578·5	803·5	478·5	67·5	2037
We have : $\dfrac{\bar{x}_s}{\sigma_x}=$	$-2·0342$	$-·9033$	$+·0936$	$+1·0837$	$+2·2293$	

leading to $$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right)=·897{,}1616, \quad \text{and} \quad r_{xC_x}=·9472.$$

If we divide into four classes, thus:

	V. R.	R.	N. H.	R. D. + D. + V. D.	Total
	109	578·5	803·5	546	2037
We have : $\dfrac{\bar{x}_s}{\sigma_x}=$	$-2·0342$	$-·9033$	$+·0936$	$+1·2253$	

leading to $$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right)=·858{,}9714, \quad \text{and} \quad r_{xC_x}=·9268.$$

And lastly if we divide into three classes, thus:

	V. R. + R.	N. H.	R. D. + D. + V. D.	Total
	687·5	803·5	546	2037
We have : $\dfrac{\bar{x}_s}{\sigma_x}=$	$-1·0826$	$+·0936$	$+1·2253$	

leading to $$S\left(\frac{n_s\bar{x}_s^2}{N\sigma_x^2}\right)=·801{,}3866, \quad \text{and} \quad r_{xC_x}=·8952.$$

Now if we put these results with those on p. 121 together we find:

Number of Groups	Gaussian Frequency Equal Subfrequencies	Pearson's Scale General Intelligence	Pearson's Scale General Health
3	·891	·881	·895
4	·928	·932	·927
5	·947	—	·947
6	·959	·966	—

It will be noticed that the frequencies differ considerably, although not extremely, from equality. Accordingly we may find it sufficient in many cases,

although the frequencies are not symmetrical or equal, to state the value of r_{xC_x} for the same number of groups with equal frequencies from the table on p. 121, and save the trouble of calculating r_{xC_x}.

Illustration III. I took the 3×3 table for correlation of health in pairs of sisters, i.e.

First Sister.

Second Sister.	V. R. + R.	N. H.	R. D. + D. + V. D.	Totals
V. R. + R. ...	428	172	87·5	687·5
N. H.	172	411	220·5	803·5
R. D. + D. + V. D.	87·5	220·5	238	546
Totals	687·5	803·5	546	2037

and I determined the contingency. The raw value of $N\phi^2 = 424{\cdot}9285$, and after correction for number of cells $= 420{\cdot}9285$. Hence

$$\phi^2 = {\cdot}206{,}6414,$$
$$C = {\cdot}4138.$$

Therefore by the value just found for r_{xC_x} for three health classes

$$r_{xy} = \frac{{\cdot}4138}{{\cdot}801{,}3866} = {\cdot}5164.$$

I then worked out r_{xy} by aid of the approximate formula

$$r_{xy} = \frac{S\left(\dfrac{n_{st}}{N}\dfrac{\bar{x}_s}{\sigma_x}\dfrac{\bar{y}_t}{\sigma_y}\right)}{S\left\{\dfrac{n_s}{N}\left(\dfrac{\bar{x}_s}{\sigma_x}\right)^2\right\} S\left\{\dfrac{n_t}{N}\left(\dfrac{\bar{y}_t}{\sigma_y}\right)^2\right\}},$$

which gives in this case $$r_{xy} = {\cdot}4940.$$

The value as given in my Huxley Lecture* for the mean of two fourfold tables with divisions first between Robust and Healthy and then between Delicate and Healthy was ·51. The agreement between the three methods appears reasonably good. I compared the last value with the same formula applied to the original 5×5 table (see *Biometrika*, Vol. III., p. 166). I found

$$S\left(\frac{n_{st}}{N}\frac{\bar{x}_s\bar{y}_t}{\sigma_x\sigma_y}\right) = {\cdot}392{,}7573^*,$$

and $$r_{xy} = {\cdot}4880.$$

It will be clear that for the type of argument which is generally based on such numbers the accordance is very satisfactory.

* In the case of the 3×3 table this product was ·317,2398, which measures the amount of correction made by denominator.

Illustration IV. I take now an extreme case, namely the Table:

A.

B.		a_1	a_2	a_3	Totals
	b_1	4983	761	49	5793
	b_2	1166	1661	572	3399
	b_3	30	248	530	808
	Totals	6179	2670	1151	10,000

I call this an extreme case because the table is only 3×3, and the frequencies and the subranges are very unequal.

We have $\quad \bar{x}_s/\sigma_x = -\cdot6172, \quad +\cdot7011, \quad +1\cdot6871,$

$\bar{y}_t/\sigma_y = -\cdot6750, \quad +\cdot7100, \quad +1\cdot8531.$

Whence I find $\quad S\left(\frac{n_s \bar{x}_s^2}{N\sigma_x^2}\right) = \cdot694{,}222, \qquad r_{xC_x} = \cdot8332,$

$$S\left(\frac{n_t \bar{y}_t^2}{N\sigma_y^2}\right) = \cdot712{,}753, \qquad r_{yC_y} = \cdot8442.$$

I then worked out the mean square contingency and found for crude numbers

$$N\phi^2 = 5146\cdot6.$$

Corrected for number of cells $\qquad \phi^2 = \cdot51426.$

Whence $\qquad C = \cdot5827.$

Therefore $\qquad r_{xy} = \cdot5827/(\cdot8332 \times \cdot8442) = \cdot8284.$

The actual table has been constructed *in round numbers* as the distribution of a Gaussian frequency surface for $r = \cdot80$, the divisions being at $x/\sigma_x = +\cdot3$ and $+1\cdot2$ and $y/\sigma_y = +\cdot2$ and $+1\cdot4$, and the correlation is hardly likely to differ by a unit from $\cdot80$. Considering the marked inequality of the subfrequencies the corrected contingency approaches closely to the correlation—at least the approach is sufficient for any argument likely to be drawn from the data in a 3×3-fold table.

Illustration V. It is profitable to show the amount of error which will be introduced by taking a marked case of non-Gaussian frequency. When looking out for such cases many years ago, I found one of the most representative instances in the case of barometric heights. I take the following table from the memoir by Dr Lee and myself*, and I have arranged the 3×3-fold table so as to give very unequal frequencies. We have thus in this case non-Gaussian frequency and very unequal frequencies. We find

$$\bar{x}_s = -1\cdot1479, \quad -\cdot1581, \quad +\cdot9161,$$

$$\bar{y}_t = -1\cdot2797, \quad -\cdot2940, \quad +\cdot8424,$$

$$S\left(\frac{n_s}{N}\frac{\bar{x}_s^2}{\sigma_x^2}\right) = \cdot768{,}494, \qquad r_{xC_x} = \cdot8823,$$

$$S\left(\frac{n_t}{N}\frac{\bar{y}_t^2}{\sigma_y^2}\right) = \cdot763{,}093, \qquad r_{yC_y} = \cdot8736.$$

* *Phil. Trans.* Vol. 190, A, p. 453.

Southampton.

Laudale.	*Over 30·15*	*30·15—29·95*	*Under 29·95*	Totals
Over *30·15* ...	545	148·5	26·5	720
30·15—29·85 ...	263·25	340·75	217·5	821·5
Under *29·85* ...	83·75	288·75	1008	1380·5
Totals ...	892	778	1252	2922

We have raw $N\phi^2 = 1449{\cdot}52$.

ϕ^2 corrected for number of cells $= {\cdot}494{,}7023$, $C = {\cdot}5753$.

Thus :
$$r_{xy} = \frac{{\cdot}5753}{{\cdot}8823 \times {\cdot}8736} = {\cdot}7464.$$

Found by the product-moment method*

Without Sheppard : $r_{xy} = {\cdot}7752 \pm {\cdot}0050$, with Sheppard : $r_{xy} = {\cdot}7802 \pm {\cdot}0049$.

Thus the difference is more than four times the probable error, but is not of a nature upon which any sweeping conclusions would be drawn. Indeed some might prefer the value drawn from the corrected 3×3-fold table, owing to the distrust of isolated outlying observations which often widely modify the constants of a distribution calculated by product-moment methods.

To still further test the pliability of the method, I rearranged the data in a 3×3-fold table with markedly unequal frequencies but nearly equal ranges, thus:

Southampton.

Laudale.	*31·05—30·15*	*30·15—29·25*	*29·25—28·35*	Totals
30·85—29·85 ...	808·25	733·25	0	1541·5
29·85—28·85 ...	83·75	1223·25	45·5	1352·5
28·85—27·85 ...	0	14	14	28
Totals ...	892	1970·5	59·5	2922

A more unpromising series of totals is hardly likely to be met with! We have

$$\bar{x}_s = -1{\cdot}1479, \quad +{\cdot}4467, \quad +2{\cdot}4133.$$

$$S\left(\frac{n_s \bar{x}_s^2}{N\sigma_x^2}\right) = {\cdot}635{,}4498, \qquad r_{xC_x} = {\cdot}7972.$$

$$\bar{y}_t = -{\cdot}7544, \quad +{\cdot}8044, \quad +2{\cdot}6797,$$

$$S\left(\frac{n_t \bar{y}_t^2}{N\sigma_y^2}\right) = {\cdot}568{,}5512, \qquad r_{yC_y} = {\cdot}7540.$$

Thus we reach some of the lowest values we have come across for r_{xC_x} and r_{yC_y} in the case of threefold divisions.

* In *Phil. Trans.* Vol. 190 A, p. 453, the value ·7572 was printed in error for ·7752.

Raw $N\phi^2 = 766{\cdot}887$, corrected $\phi^2 = {\cdot}761{,}0838$ and $C = {\cdot}4550$,

$$r_{xy} = {\cdot}4550/({\cdot}7972 \times {\cdot}7540) = {\cdot}7570,$$

a result differing only by ·02 from that of the product-moment method.

It will be noticed that I have arranged the material so that the subranges are *equal*. If we work the 3×3-fold table out, using Sheppard's correction, we find by product-moment

$$r_{xy} = {\cdot}7746,$$

which is considerably nearer the mark than the corrected contingency value.

In this case a poor result is given by the formula:

$$S\left(\frac{n_{st}}{N}\frac{\bar{x}_s}{\sigma_x}\frac{\bar{y}_t}{\sigma_y}\right) \Big/ S\left(\frac{n_s \bar{x}_s^2}{N\sigma_x^2}\right) \times S\left(\frac{n_t \bar{y}_t^2}{N\sigma_y^2}\right),$$

the assumption that $S(n_{st}\bar{x}_{st}\bar{y}_{st}) = S(n_{st}\bar{x}_s\bar{y}_t)$ being by no means satisfactory for (i) very non-Gaussian distributions, or (ii) high correlations, when the subranges if equal are very large. For *equal* ranges with a moderate approach to the Gaussian distribution and a correlation not much exceeding 0·5, it gives quite good results. Thus if we take the table for stature in Father and Son and arrange it in practically equal subranges thus:

Stature of Father.

Stature of Son.	*58·5—64·5*	*64·5—70·5*	*70·5—76·5*	Totals
Under *67·5* ...	86·25	269	9·75	365
67·5—73·5 ...	39·25	495·5	129·25	664
73·5—79·5 ...	1	22	26	49
Totals ...	126·5	786·5	165	1078

we find:

$$\bar{x}_s = -1{\cdot}6780, \quad -{\cdot}0540, \quad +1{\cdot}5439,$$

$$\bar{y}_t = -1{\cdot}0804, \quad +{\cdot}4387, \quad +2{\cdot}1022,$$

$$S\left(\frac{n_s \bar{x}_s^2}{N\sigma_x^2}\right) = {\cdot}697{,}738, \qquad r_{xC_x} = {\cdot}8353,$$

$$S\left(\frac{n_t \bar{y}_t^2}{N\sigma_y^2}\right) = {\cdot}714{,}644, \qquad r_{yC_y} = {\cdot}8454,$$

$$S\left(\frac{n_{st}\bar{x}_s\bar{y}_t}{N\sigma_x\sigma_y}\right) = {\cdot}260{,}7392,$$

$$\therefore \quad r_{xy} = {\cdot}5229,$$

against ·5259 by the product-moment methods.

The same table yields by 3×3-fold contingency

$$N\phi^2 = 163{\cdot}909, \qquad \text{corrected } \phi^2 = {\cdot}148{,}3386, \qquad C = {\cdot}3594,$$

whence $r_{xy} = {\cdot}5089$.

Thus the corrected contingency for roughly equal frequencies and equal sub-ranges gives respectively ·52 and ·51, but the formula (xvii) gives for equal ranges ·52 and for roughly equal frequencies ·57. It seems to me therefore that corrected contingency can be more safely applied than formula (xvii).

Conclusions.

A further discussion of the corrections needful when using the method of mean square contingency will shortly be published, but the present paper seems to indicate that it can in the bulk of cases for 3×3-fold, 4×3-fold or 4×4-fold tables be used effectively. I am not aware that any other effective method has been proposed for such tables. The assumption of the Gaussian distribution to determine the correlation of the variate with its class index need not be made, if the material thrown into broad classes has had a sample quantitatively determined. But we have seen in this paper that the Gaussian assumption to fix the means of the broad classes is in many cases amply sufficient to give good results even in non-Gaussian frequencies. Of course a control series to determine r_{xC_x}, where practicable, may be advantageously sought*, but until some better method can be suggested, the present seems to me the best available for dealing with the correlation of variates classed in a few broad categories.

With contingency tables of 5×5 or 6×6, the correcting factors will generally be small, but corrections must certainly be made for 4×4 and 3×3 tables.

* For example, I took Southampton frequency, and calculated from the original table the means of the three groups 892, 1970·5, 59·5, corresponding to the table on p. 137; they were $-1·1006$, $+·4123$, and $+2·1746$, leading to $r_{xC_x} = ·8056$, as against the Gaussian value ·7972. This is for a sensibly non-Gaussian frequency, and a much worse agreement is permissible in actual practice, where the nature of the argument rarely turns on correlation differences of less than ·05.

BIBLIOGRAPHY*.

(1) Albrand, W. Über das Ergebnis von Augenuntersuchungen innerhalb der niedersächsischen Rasse und seine Bedeutung für einige rassenbiologische Beziehungen. Arch. f. Rassen- u. Gesells. Biol. 1911, Bd. viii, S. 33—58. Leipzig, 1911.

(2) Alexander, W. B. Further Experiments on the Cross-breeding of Two Races of the Moth *Acidalia virgularia*. Proc. Roy. Soc., B, Vol. lxxxv, pp. 45—52. London, 1912.

(3) Assheton, R. The geometrical relation of the nuclei in an invaginating gastrula (e.g. *Amphioxus*) considered in connection with cell rhythm and Driesch's conception of Entelechy. Arch. f. Entwicklungsmech. der Organis., Bd. xxix, S. 46—78. Leipzig, 1910.

(4) Bashford, E. F. The Incidence of Cancer in Mice of Known Age. Proc. Roy. Soc., B, Vol. lxxxi, pp. 310—313. London, 1909.

(5) Bateson, W. and Punnett, R. C. On the Interrelations of Genetic Factors. Proc. Roy. Soc., B, Vol. lxxxiv, pp. 3—8. London, 1911.

(6) Bateson, W. Biological Fact and the Structure of Society. The Herbert Spencer Lecture, Oxford, 1912.

(7) Bauer, E. Propfbastarde. Biol. Centralbl., Bd. xxx, S. 497—514. Leipzig, 1910.

(8) Baur, E. Einführung in die experimentelle Vererbungslehre. Bornträger. Berlin, 1911.

(9) Bayerthal. Über den gegenwärtigen Stand meiner Untersuchungen über die Beziehungen zwischen Kopfgrösse und Intelligenz im schulpflichtigen Alter. Internat. Arch. f. Schulhygiene, Bd. vii, S. 244—262. Munich, 1910.

(10) Bean, R. B. Filipino Types. I. Manila Students. An attempt to classify the littoral population of Luzon and adjacent islands. II. Types found in Malecon Morgue. Philippine Jour. of Sci., Sect. A, Vol. xiv, pp. 263—337. Manila, 1909.

(11) Becher, E. Theoretische Beiträge zum Darwinismus. Arch. f. Rassen- u. Gesells. Biol. Bd. vii, S. 137—158, 265—289. Leipzig, 1910.

(12) Beneduce, A. Le Variazioni di Mortalità secondo gli Anni di Età. Giorn. d. Econ., Ser. 2, Anno xviii, pp. 1069—1104. Rome, 1907.

(13) Benini, R. Una Possibile Creazione del Metodo Statistico. (L'Economia politica induttiva.) Giorn. d. Econ., Ser. 2, Anno xix, pp. 11—34. Rome, 1908.

(14) Berry, R. J. A., Robertson, A. W. D. and Stuart Cross, K. A Biometrical Study of the Relative Degree of Purity of Race of the Tasmanian, Australian and Papuan. Proc. Roy. Soc. of Edin., Vol. xxxi, pp. 17—40. Edinburgh, 1910–11.

(15) Berry, R. J. A. and Robertson, A. W. D. The Place in Nature of the Tasmanian Aboriginal as deduced from a Study of his Calvarium. Part I. His Relations to the Anthropoid Apes, Pithecanthropus, Homo primigenius, Homo fossilis and Homo sapiens. Proc. Roy. Soc. of Edin., Vol. xxxi, pp. 41—69. Edinburgh, 1910–11.

* This bibliography includes all papers sent to this Journal. Authors will oblige the Editor by adding volume, pagination and place of publication of the original transactions, proceedings or journals, when the paper is an offprint.

(16) Bertillon, J. La dépopulation de la France, ses conséquences, ses causes, mesures à prendre pour la combattre. F. Alcan, Paris, 1911.

(17) Betz, W. Untersuchungen von F. Galton, K. Pearson und ihrer Schule über Begabung und Vererbung. Zeitschr. f. angewandte Psychologie, Bd. III, S. 273—280. Leipzig, 1909.

(18) ——. Über Korrelation. Methoden der Korrelationsberechnung und kritischer Berichte über Korrelationsuntersuchungen aus dem Gebiete der Intelligenz, der Anlagen und ihrer Beeinflussung durch äussere Umstände. Beihefte z. Zeitschr. f. angewandte Psychologie, Heft 3. Leipzig, 1911.

(19) Biasutti, R. Contributi all' antropologia e all' antropogeografia delle popolazioni del Pacifico settentrionale. Arch. per l'Antrop. e l'Etnol., T. XL, p. 82. Florence, 1910.

(20) Boas, F. Psychological Problems in Anthropology. Amer. Jour. of Psych., Vol. XXI, pp. 371—384. Worcester, Mass., 1910.

(21) Bolk, L. Über die Obliteration der Nähte am Affenschädel, zugleich ein Beitrag zur Kenntnis der Nahtanomalien. Zeitschr. f. Morph. u. Anthrop., Bd. XV. S. 1—206. Stuttgart, 1912.

(22) Bresciani, C. Sui Metodi per la Misura delle Correlazione. Giorn. d. Econ., Ser. 2, Anno XX, pp. 401—414, 491—516. Rome, 1909.

(23) Brown, W. B. The Use of the Theory of Correlation in Psychology. Cambridge, 1910.

(24) ——. Some Experimental Results in the Correlation of Mental Abilities. Brit. Jour. of Psych., Vol. III, pp. 296—322. Cambridge, 1910.

(25) ——. Educational Psychology in the Secondary Schools. Jour. of Philosophy, Psychology and Scientific Methods, New York, Vol. VII, pp. 14—18, 1910.

(26) Bruyker, C. de. Over dubbele halve curven. Handel. v. h. twaalfde Vlaamsch Nat. en Geneesk. Cong., pp. 215—224. St Niklaas, 1908.

(27) ——. De heterostylie bij Primula eliator Jacq., statistiche gegevens. Handel. v. h. twaalfde Vlaamsch Nat. en Geneesk. Cong., pp. 241—248. St Niklaas, 1908.

(28) ——. De statistische Methode in de Plantkunde en hare toepassingen op de studie van den Invloed der Levensvoorwaarden. A. Siffer, Ghent, 1910.

(29) Buchanan, Florence. The Significance of the Pulse Rate in Vertebrate Animals. Smithsonian Report for 1910, Washington, pp. 487—505. 1911.

Contains several tables giving averages of weight, carbon-dioxide output, heart weight in percentage of body weight, and average frequency of heart beat for 12 different birds varying in size from a Goldfinch to a Goose, and also for 15 mammals varying in weight from a Mouse to a Race Horse. The author finds "a reciprocal relation between pulse rate and relative heart size dependent upon the rate at which oxidation processes occur." The method of taking the pulse of birds and small animals is explained and illustrated.

(30) Bulloch, W. The Problem of Pulmonary Tuberculosis considered from the Standpoint of Infection. Horace Dobell Lecture, London, Nov. 1910.

Contains a very interesting historical sketch of the subject and gives data concerning the prevalence of the disease and the possibility of its arising by means of infected milk in Germany, Japan, Norway, Turkey, Rumania, Egypt, Gold Coast, Asiatic Turkey, Singapore, Philippine Islands, New Zealand, Faroe Islands and Greenland.

(31) Bulloch, W. and Greenwood, M. The Problem of Pulmonary Tuberculosis considered from the Standpoint of Disposition. Proc. Roy. Soc. of Med., Vol. IV, Pt. 2, pp. 147—184. London, 1911.

(32) Calhoun, L. A. Sex Determination and its Practical Application. Eugenics Pub. Co., New York City, 1910.

(33) Cannon, W. A. Studies in Heredity as illustrated by the Trichomes of Species and Hybrids of Juglans, Œnothera, Papaver and Solanum. Carnegie Inst. of Washington Pub. No. 117. 1909.

(34) Cantacuzene, G. Contribution à la Craniologie des Romains Anciens. L'Anthropologie, T. xxi, pp. 55—74. Paris, 1910.

(35) Carpena, F. Antropologia Criminal. Libreria de Fernando Fé, Madrid, 1909.

(36) Castle, W. E. and Phillips, J. C. On Germinal Transplantation in Vertebrates. Carnegie Inst. of Washington Pub. No. 144. 1911.

(37) Castle, W. E. The Inconstancy of Unit Characters. Amer. Naturalist, Vol. xlvi, pp. 352—362. New York, 1912.

(38) Chartier, C. V. L. Contributions to the Mathematical Theory of Statistics.

I. Theorems of Poisson and Lexis. Arkiv f. Matematik, Astronomi och Fysik, Stockholm, Bd. vii, No. 17. 1911.

II. Homograde and heterograde statistics. *Ibid.* Bd. viii, No. 2. 1912.

III. Some fundamental problems in the theory of probability. *Ibid.* Bd. viii, No. 4. 1912.

IV. The statistical series in homograde statistics. *Ibid.* Bd. viii, No. 4. 1912.

(39) Chisholm, R. A. On the size and growth of the blood in tame rats. Quart. Jour. of Exper. Physiol., Vol. iv, pp. 207—229. London, 1911.

(40) Christophers, S. R. Suggestions on the use of Available Statistics for studying Malaria in India. Paludism, No. i, pp. 16—32. Simla, 1910.

(41) ——. Malariometry. Observations upon Graphs of the Spleen Rate and the Average Spleen. Paludism, No. iii, pp. 87—102. Simla, 1911.

(42) Cillis, E. De. Intorno ad un possibile nuovo metodo di determinazione delle razze di piante coltivate con speciale riguardo al frumento. Atti del R. Istituto d'Incoraggiamento di Napoli, Ser. vi, T. ix, pp. 173—198. Naples, 1911.

(43) Claasen, W. Beiträge zur Feststellung der Ernährungsverhältnisse des deutschen Land- und Stadtvolkes. Arch. f. Rassen- u. Gesells. Biol., Bd. viii, 1911, S. 458—487, 604—627. Leipzig, 1911.

(44) Clark, C. F. Variation and Correlation in Timothy. Cornell Univ. Agr. Exp. Stat. Bull. 279, pp. 301—349. Ithaca, N.Y., 1910.

(45) Craig, J. I. An Anthropometrical Survey of Egypt. The Cairo Scientific Journal, Vol. v, pp. 165—180. Cairo, 1911.

The data for this paper were obtained from measurements of head length, head breadth, length of left middle finger, length of left foot, left cubit, and stature taken on 10,000 Egyptian prisoners.

(46) Cross, K. Stuart. On a Numerical Determination of the Relative Positions of certain Biological Types in the Evolutionary Scale, and of the Relative Values of various Cranial Measurements and Indices as Criteria. Proc. Roy. Soc. of Edin., Vol. xxxi, pp. 70—84. Edinburgh, 1910–11.

(47) Cunningham, J. T. Mendelian Experiments on Fowls. Proc. Zool. Soc., Vol. i, pp. 241—259. London, 1912.

(48) Curtis, M. R. An accurate method for determining the weight of the parts of the eggs of birds. Ann. Rep. of Maine Agr. Exp. Stat., pp. 93—112. Orono, Maine, 1911.

(49) Darwin, L. Sociology and Evolution. Charity Organization Review, N. S., Vol. xxvi, pp. 364—378. London, 1909.

(50) DAUDIN. Travaux et problèmes relatifs à la Parthénogénèse artificielle. Bulletin Scientifique de la France et de la Belgique, T. XLIII, pp. 299—372. Paris, 1909.

(51) DAVENPORT, GERTRUDE C. and C. B. Heredity of Skin Pigmentation in Man. Amer. Naturalist, Vol. XLIV, pp. 641—672, 705—731. New York, 1910.

(52) DAVENPORT, C. B. Inheritance of plumage color in poultry. Proc. Soc. for Exper. Biol. and Med., Vol. VII, p. 168. New York, 1910.

(53) ——. Fit and Unfit Matings. The Training School, Vol. VII, pp. 258—262. New Jersey, 1910.

(54) ——. Euthenics and Eugenics. Pop. Sci. Mon., pp. 16—20. New York, 1911.

(55) ——. The Transplantation of Ovaries in Chickens. Jour. of Morph., Vol. XXII, pp. 111—122. Philadelphia, 1911.

(56) DAVENPORT, C. B. and WEEKS, D. F. A first study of inheritance in epilepsy. Jour. of Mental Disease, Vol. XXXVIII, pp. 641—670. New York, 1911.

(57) DENDY, A. Momentum in Evolution. Brit. Assoc., Portsmouth Meeting, Vol. LXXXI, pp. 277—280. 1911.

(58) DONALDSON, H. H. Further observations on the nervous system of the American Leopard Frog (*Rana pipiens*) compared with that of the European Frogs (*Rana esculenta* and *Rana temporaria*). Jour. of Comp. Neurol. and Psych., Vol. XX, pp. 1—18. Philadelphia, 1910.

(59) ——. On the Percentage of Water in the Brain and in the Spinal Cord of the Albino Rat. Jour. of Comp. Neurol. and Psych., Vol. XX, pp. 119—144. Philadelphia, 1910.

(60) ——. On the Regular Seasonal Changes in the Relative Weight of the Central Nervous System of the Leopard Frog. Jour. of Morph., Vol. XXII, pp. 663—694. Philadelphia, 1911.

(61) ——. On the Influence of Exercise on the Weight of the Central Nervous System of the Albino Rat. Jour. of Comp. Neurol. and Psych., Vol. XXI, pp. 129—137. Philadelphia, 1911.

(62) ——. The Effect of Underfeeding on the Percentage of Water, on the Ether-Alcohol Extract, and on Medullation in the Central Nervous System of the Albino Rat. Jour. of Comp. Neurol. and Psych., Vol. XXI, pp. 139—145. Philadelphia, 1911.

(63) DONALDSON, H. H. and HATAI, S. Note on the Influence of Castration on the Weight of the Brain and Spinal Cord in the Albino Rat and on the Percentage of Water in them. Jour. of Comp. Neurol. and Psych., Vol. XXI, pp. 155—160. Philadelphia, 1911.

(64) DONALDSON, H. H. An Interpretation of some differences in the Percentage of Water found in the Central Nervous System of the Albino Rat and due to conditions other than age. Jour. of Comp. Neurol. and Psych., Vol. XXI, pp. 161—176. Philadelphia, 1911.

(65) DONALDSON, H. H. and HATAI, S. A comparison of the Norway Rat with the Albino Rat in respect to Body Length, Brain Weight, Spinal Cord Weight and the Percentage of Water in both the Brain and the Spinal Cord. Jour. of Comp. Neurol. and Psych., Vol. XXI, pp. 417—458. Philadelphia, 1911.

(66) DONALDSON, H. H. A comparison of the European Norway and Albino Rats (*Mus Norvegicus* and *Mus Norvegicus Albinus*) with those of North America in respect to the weight of the Central Nervous System and to Cranial Capacity. Jour. of Comp. Neurol. and Psych., Vol. XXII, pp. 71—97. Philadelphia, 1912.

The author compares Norway and Albino rats from Paris, London and Vienna with those of Philadelphia: he concludes that they are essentially similar in respect to the characters considered. The samples measured are small, particularly so in the case of the albinos.

(67) Donaldson, H. H. On the weight of the crania of Norway and Albino Rats from three stations in Western Europe and one station in the United States. Anat. Record, Vol. vi, pp. 53—63. Philadelphia, 1912.

(68) Doncaster, L. and Marshall, F. H. A. The effects of one-sided ovariotomy on the sex of the offspring. Jour. of Genetics, Vol. i, pp. 70—72. Cambridge, 1910.

(69) Doncaster, L. Recent work on the determination of sex. Sci. Prog., pp. 90—104. London, 1909.

(70) Driesmans, H. Menschenreform und Bodenreform. Unter Zugrundelegung der Veredelungslehre Francis Galtons. Dietrich, Leipzig, 1911.

(71) Drinkwater, H. The Inheritance of "Thumb-fingerness" in a Short-fingered Family. Mendel Journal, pp. 35—52. London, 1911.

(72) ——. Account of a Family showing minor-Brachydactyly. Jour. of Genetics, Vol. ii, pp. 21—40. 1912.

(73) East, E. M. The Distinction between Development and Heredity in Inbreeding. Amer. Naturalist, Vol. 43, pp. 173—181. New York, 1909.

(74) ——. A Mendelian Interpretation of Variation which is apparently continuous. Amer. Naturalist, Vol. xliv, pp. 65—82. New York, 1910.

(75) ——. Inheritance in Potatoes. Amer. Naturalist, Vol. xliv, pp. 424—430. New York, 1910.

(76) Eccles, R. G. Parasitism and Natural Selection. A medical supplement to Darwin's Origin of Species. Medical Record, Vol. lxxvi, pp. 169—177, 757—765. New York, 1909.

(77) ——. Importance of disease in plant and animal evolution. Medical Record, Vol. lxxxi, pp. 501—509. New York, 1912.

(78) Edgeworth, F. Y. Applications of Probabilities to Economics. The Economic Journal, Vol. xx, pp. 284—465. London, 1910.

(79) ——. On the Application of the Calculus of Probabilities to Statistics. Internat. Stat. Inst., Bull. No. 18. Paris, 1910.

(80) Ehrenfels, C. von. Leitziele zur Rassenbewertung. Arch. f. Rassen- u. Gesells. Biol. Bd. viii, S. 59—71. Leipzig, 1911.

(81) Elderton, E. M. On the marriage of first cousins. Galton Eugenics Laboratory Lecture Series, No. iv, Dulau & Co. Ltd. 1911.

(82) Elderton, W. P. and Perry, S. J. A Third Study of the Statistics of Pulmonary Tuberculosis. The Mortality of the Tuberculous and Sanatorium Treatment. Biometric Laboratory Publications, Studies in National Deterioration, vi, Dulau & Co. Ltd. London, 1910.

(83) Eller, R. Familiärer Kretinismus in Wien. Jahrb. f. Kinderheilkunde u. phys. Erziehung, Bd. lxxi, S. 586—760. Berlin, 1910.

(84) Ewing, E. C. Correlation of Characters in Corn. Cornell Univ. Agr. Exp. Stat. Bull. 287, pp. 67—100. Ithaca, N.Y., 1910.

(85) Federley, H. Vererbungsstudien an der Lepidopteren-Gattung Pygaera. Arch. f. Rassen- u. Gesells. Biol., Bd. viii, S. 281—337. Leipzig, 1911.

(86) Fehlinger, H. De l'Influence Biologique de la Civilisation Urbane. Scientia, Bologna, T. x, pp. 421—434. 1911.

(87) ——. Ist Alkoholismus eine Ursache der Entartung? Arch. f. Kriminalanthrop. u. Kriminalistik, Leipzig, Bd. xli, S. 302—306. Berlin, 1911.

(88) FEHLINGER, H. Über einige Faktoren der progressiven biologischen Entwicklung der Kulturvölker. Dokumente des Fortschritts, S. 224—228. Berlin, 1912.

(89) FENTON, F. The Influence of Newspaper Presentations upon the Growth of Crime and other Anti-Social Activity. Amer. Jour. of Sociology, Vol. XVI, pp. 342—371, 538—564. Chicago, 1910 and 1911.

(90) FERRÁNDIZ, M. A. Y. Los Origenes Etnicos de las Nacionalidades Libio-Ibéricas. Ciencias Naturales, Madrid, Sección 4ª. 1910.

(91) FIELD, J. A. The Early Propagandist movement in English Population Theory. Amer. Economic Review, April, 1911, and Bull. Amer. Econ. Assoc. 23rd Annual Meeting, pp. 207—236, Princeton, N.J.

(92) ——. The Progress of Eugenics. Quart. Jour. of Economics, Vol. XXVI, pp. 1—67. Cambridge, U.S.A., 1911.

(93) FIELD, J. H. Statement of actual rainfall in June, July, August and September, 1911, and a comparison of the forecasts with the actual rainfall. Calcutta, 1911.

(94) ——. Memorandum on the monsoon conditions prevailing during June and July with a forecast for August and September. Simla, 1911.

(95) ——. Forecast of the probable character of the rains during December 1911 and January and February 1912 in Northern India. Calcutta, 1911.

(96) FISCHER, E. Ein Fall von erblicher Haararmut und die Art ihrer Vererbung. Arch. f. Rassen- und Gesells. Biol., Bd. VII, S. 50—55. Leipzig, 1910.

(97) ——. Zur Frage der "Kreuzungen beim Menschen." Arch. f. Rassen- u. Gesells. Biol., Bd. IX, S. 8—9. Leipzig, 1911.

(98) FISHER, R. A. On an Absolute Criterion for Fitting Frequency Curves. Messenger of Mathematics, New Series, Vol. XLI, pp. 155—160. Cambridge, 1912.

(99) FRANZ, S. I. F. and WHITE, W. A. The Use of Association Tests in determining mental Contents. Gov. Hosp. for Insane, Washington, Bull. No. 1. 1909.

(100) FRIEDENTHAL, H. Beiträge zur Naturgeschichte des Menschen. V. Sonderformen der menschlichen Leibesbildung. G. Fischer, Jena, 1910.

(101) FURLAN, V. Note sulla Curva Paretiana dei Redditi. Giorn. degli Econ., Ser. 2, Anno XX, pp. 695—701. Rome, 1909.

(102) GALLARDO, A. Recientes Contribuciones Matematicas al estudio de las Leyes de la Herencia Biologica. Anales de la Soc. Cientifica Argentina, T. LXVIII, pp. 185 et seq. Buenos Aires, 1909.

(103) GARNETT, W. On representation of certain examination results in two and in three dimensions. Jour. Roy. Stat. Soc., Vol. LXXIII, pp. 1—9. London, 1910.

(104) GATES, R. R. Studies of Inheritance in the Evening Primrose. Chicago Medical Recorder. Chicago, Feb. 1909.

(105) ——. Apogamy in Oenothera. Science, N. S., Vol. XXX, pp. 691—694. New York, 1909.

(106) ——. The Behavior of Chromosomes in Oenothera Lata × O. Gigas. Bot. Gazette, Vol. XLVIII, pp. 179—197. Chicago, 1909.

(107) ——. Abnormalities in Oenothera. 21st Ann. Rep. of Missouri Bot. Garden, pp. 175—184. St Louis, 1910.

(108) ——. The Material Basis of Mendelian Phenomena. Amer. Naturalist, Vol. XLIV, pp. 203—213. New York, 1910.

(109) ——. Mutation in Oenothera. Amer. Naturalist, Vol. XLV, pp. 577—606. New York, 1911.

(110) Gates, R. R. Pollen Formation in Oenothera Gigas. Annals of Botany, Vol. xxv, pp. 909—940. Oxford, 1911.

(111) ——. Studies on the Variability and Heritability of Pigmentation in Oenothera. Zeitsch. f. indukt. Abstam.- und Vererbungslehre, Bd. iv, S. 337—372. Berlin, 1911.

(112) ——. The Mode of Chromosome Reduction. Bot. Gaz., Vol. li, pp. 321—344. Chicago, 1911.

(113) Gerould, J. H. The Inheritance of Polymorphism and Sex in Colias philodice. Amer. Naturalist, Vol. xlv, pp. 257—283. New York, 1911.

(114) Giuffrida-Ruggeri, V. Alcune Idee Controverse sul Dimorfismo Sessuale nell' Uomo. Arch. per l'Antrop. e l'Etnol., T. xl, pp. 2—8. Florence, 1910.

(115) ——. La Quistione dei Pigmei e le variazioni morfologiche dei Gruppi etnici. Arch. per l'Antrop. e l'Etnol., T. xl, fasc. 3—4. Florence, 1910.

(116) ——. L'Uomo comè specie collettiva. Naples, 1912.

(117) Glynn, E. E. and Cox, G. L. Variations in the Inherent Phagocytic Power of Leucocytes. Jour. of Pathology and Bacteriology, Vol. xiv, pp. 90—131. Edinburgh, 1909.

(118) Goddard, H. H. Heredity of Feeble-Mindedness. Amer. Breeder's Mag., Vol. i, pp. 165—178. Washington, 1910.

Contains 19 pedigrees of mentally defective stocks.

(119) Goldschmidt, R. Das Problem des Geschlechtsbestimmung. Die Umschau, March, 1910.

(120) Goodale, H. D. Breeding Experiments in Poultry. Proc. Soc. for Exp. Biol. and Med., Vol. vii, pp. 178—179. New York, 1910.

(121) Gordon, W. The Influence of Soil on Phthisis as illustrating a neglected Principle in Climatology. Brit. Med. Jour., Sept. 25th, 1909.

(122) ——. Prevalent rain-bearing winds and the milk supply of Manchester during the years 1896-1909. Brit. Med. Jour., 1911, Vol. i, pp. 482—485. London, 1911.

(123) Goring, C. On the Inheritance of the Diatheses of Phthisis and Insanity. A Statistical Study based upon the Family History of 1500 Criminals. Biometric Laboratory Publications, Studies in National Deterioration, v, Dulau & Co., Ltd. London, 1909.

(124) Gortner, R. A. Spiegler's "White Melanin" as related to Dominant or Recessive White. Amer. Naturalist, Vol. xliv, pp. 497—502. New York, 1910.

(125) ——. On Melanin. Biochemical Bulletin, Vol. i, pp. 207—215. Cold Spring Harbor, N.Y., 1911.

(126) ——. Studies on Melanin. iii. The Inhibition by certain Phenolic Substances of the Action of Tyrosinase. (A suggestion as to the cause of Dominant and Recessive Whites.) Jour. Biol. Chem., Vol. x, p. 113. New York, 1911.

(127) ——. Studies on Melanin. iv. The origin of the pigment and the color pattern in the Elytra of the Colorado Potato Beetle. Amer. Naturalist, Vol. xlv, pp. 743—755. New York, 1911.

(128) Greenwood, M. The Problem of Marital Infection in Pulmonary Tuberculosis. Proc. Roy. Soc. of Med., Vol. ii, Pt. 2, Epid. Section, pp. 259—282. London, 1909.

(129) Greenwood, M., and Candy, R. H. The Fatality of Fractures of the Lower Extremity and of Lobar Pneumonia. A Study of Hospital Mortality Rates, 1751-1901. Jour. Roy. Stat. Soc., Vol. lxxiv, pp. 365—405. London, 1911.

(130) Greenwood, M. Statistical Investigation of Plague in the Punjab. Second Report: on the connection between Proximity to Railways and Frequency of Epidemics. Jour. of Hygiene, Vol. xi, Supplement, pp. 49—61. Cambridge, 1911.

(131) GREENWOOD, M. Statistical Investigation of Plague in the Punjab: on some of the Factors which influence the Prevalence of Plague. Jour. of Hygiene, Vol. XI, Supplement, pp. 62—156. Cambridge, 1911.

(132) GROBER. Die Behandlung der Rassenschäden. Arch. f. Rassen- u. Gesells. Biol., Bd. IX, S. 49—86. Leipzig, 1912.

(133) GROSS, J. Über Vererbung und Artbildung. Biol. Centralbl., Bd. XXXI, S. 161—214. Leipzig, 1911.

(134) GRUBER, M. Vererbung, Auslese und Hygiene. Deutsch. Med. Wochenschrift. 35 Jahrg., S. 1993—1996, 2049—2053. Berlin, 1909.

(135) GURLEY, R. R. Chapters for a Biological-Empirical Psychology. Amer. Jour. of Psych., Vol. XX, pp. 60—106. Worcester, Mass., 1909.

(136) GUYER, M. F. Atavism in Guinea-Chicken Hybrids. Jour. of Exp. Zool., Vol. VII pp. 723—745. Baltimore, 1909.

(137) HÄCKER, V. Allgemeine Vererbungslehre. Fr. Vieweg u. Sohn, Brunswick, 1911.

(138) HAGEDOORN, A. L. Mendelian inheritance of sex. Arch. f. Entwick. Mech., Bd. XXVIII, S. 1—34. Leipzig, 1909.

(139) ——. Autokatalytical substances, the determinants for the inheritable characters. A biomechanical theory of inheritance and evolution. W. Engelmann, Leipzig, 1911.

(140) HALE-WHITE, W. The Outlook of Sufferers from Exophthalmic Goitre. Quart. Jour. of Med., Vol. IV, pp. 89—108. Oxford, 1910.

(141) HARRIS, J. A. Is there a Selective Elimination of Ovaries in the Fruiting of the Leguminosae? Amer. Naturalist, Vol. XLIII, pp. 556—559. New York, 1909.

(142) ——. Correlation between Length of Flowering, Stalk and Number of Flowers per Inflorescence in Nothoscordum and Allium. 20th Ann. Rep. of Missouri Bot. Garden, pp. 105—115. St Louis, 1909.

(143) ——. Correlations in the Inflorescence of Celastrus Scandens. 20th Ann. Rep. of Missouri Bot. Garden, pp. 116—122. St Louis, 1909.

(144) ——. The Leaves of Podophyllum. Bot. Gaz., Vol. XLVII, pp. 438—444. Chicago, 1909.

(145) ——. A Bimodal Variation Polygon in Syndesmon Thalictroides and its Morphological Significance. Amer. Naturalist, Vol. XLIV, pp. 19—30. New York, 1910.

(146) ——. The Arithmetic of the Product Moment Method of Calculating the Coefficient of Correlation. Amer. Naturalist, Vol. XLIV, pp. 693—699. New York, 1910.

(147) ——. On the Relationship between the Length of the Pod and Fertility and Fecundity in Cercis. Bot. Gaz., Vol. L, pp. 117—127. Chicago, 1910.

(148) ——. Correlation in the Inflorescence of Sanguinaria. Biol. Centralb., Vol. XXX, pp. 629—633. Leipzig, 1910.

(149) ——. A First Study of the Influence of Starvation of the Ascendants upon the Characteristics of Descendants. Amer. Naturalist, Vol. XLVI, pp. 313—343, and pp. 656—674. New York, 1912.

(150) HARTOG, P. J. Examinations in their Bearing on National Efficiency. Jour. of Roy. Soc. of Arts, London, Vol. LIX, pp. 279—288. 1911.

(151) HARVEY, H. W. The Action of Poisons upon Chlamydomonas and other Vegetable Cells. Annals of Botany, Vol. XXIII, pp. 181—187. Oxford, 1909.

(152) HATAI, S. De Forest's Formula for "An Unsymmetrical Probability Curve." Anat. Rec., Vol. IV, pp. 281—290. Philadelphia, 1910.

(153) ——. On the length of the internodes in the sciatic nerve of Rana temporaria (fusca) and Rana pipiens. Being a reexamination by biometric methods of the data studied by Boycot, 1904, and Takahashi, 1908. Jour. of Comp. Neurol. and Psych., Philadelphia, 1910, pp. 19—47.

(154) HATAI, S. A Mathematical Treatment of some Biological Problems. Biol. Bull., Vol. XVIII, pp. 126—130. Philadelphia, 1910.

(155) ——. The Mendelian Ratio and Blended Inheritance. Amer. Naturalist, Vol. XLV, pp. 99—106. New York, 1911.

(156) ——. A Formula for Determining the Total Length of the Leopard Frog (*R. Pipiens*) for a Given Body Weight. Anat. Record, Vol. V, pp. 309—312. Philadelphia, 1911.

(157) ——. An Interpretation of Growth Curves from a Dynamical Standpoint. Anat. Record, Vol. V, pp. 373—382. Philadelphia, 1911.

(158) HANSCHILD, M. W. Untersuchung über die Pigmentation in Auge verschiedener Menschenrassen und die Pigmentation in Säugetierauge überhaupt. Zeitschr. f. Morph. u. Anthrop., Stuttgart, Bd. XII, S. 473—544. 1910.

(159) HEGNER, R. W. The origin and early history of the germ-cells in some Chrysomalid beetles. Jour. of Morph., Vol. XX, pp. 231—296. Philadelphia, 1909.

(160) HELGUERO, F. DE. Sulla Rappresentazione Analitica delle curve statistiche. Giorn. degli Econ., Ser. 2, Anno XX, pp. 241—265. Rome, 1909.

(161) HEMMETER, J. C. Anthropometric studies of the osseous Proportions of the human Body, with a view to obtaining a mathematic expression for Enteroptosis. Internat. Beiträge z. Pathol. u. Therapie der Ernährungsstörungen, Berlin, Bd. II, Heft 3. 1910.

(162) HERON, D. The influence of defective physique and unfavourable home environment on the intelligence of school children, being a statistical examination of the London County Council pioneer School Survey. Galton Eugenics Laboratory Memoir Series, VIII. Dulau & Co., Ltd. 1910.

(163) HEYER, A. Neue Untersuchungen über die Längenvariation der Koniferennadeln. Berichte der Schweizerischen botanischen Gesellschaft, Heft XX, S. 247—253. Berne, 1911.

(164) HILL, J. A. Studies on Strength and Elasticity of the Wool Fiber. I. The Probable Error of the Mean. Paper from the Wool Laboratory of the Univ. of Wyoming Experimental Station. No. 1. Supplement to Twenty-first Annual Report.

The problem discussed is the following: How many fibres must be tested in a given sample of wool in order that the average breaking strain and elasticity found for the tested fibres may represent satisfactorily the average for the whole sample?

(165) HOGGAN, FRANCES. The American Negro and Race Blending. Sociological Review, Vol. II, pp. 349—360. London, 1909. (A very inadequate discussion of a great problem.)

(166) HORT, E. C. and PENFOLD, W. J. A Critical Study of Experimental Fever. Proc. Roy. Soc., B, Vol. LXXXV, pp. 174—186. London, 1912.

(167) HRDLICKA, A. Note sur la variation morphologique des Égyptiens depuis les temps préhistoriques ou prédynastiques. Bull. de la Soc. d'Anthrop. de Paris, Série V, T. X, pp. 143—144. Paris, 1909.

(168) ——. Contribution to the Anthropology of Central and Smith Sound Eskimo. Anthrop. Papers of Amer. Mus. of Nat. Hist., Vol. V, pp. 177—280. New York, 1910.

(169) ——. Human Dentition and Teeth from the Evolutionary and Racial Standpoint. The Dominion Dental Journal. 1911.

The teeth of mankind to-day are compared with those of the earliest identified human lower jaw, that of the "man of Heidelberg"; also with those of the rather less ancient jaws of one of the cave men of ancient France and of the "man of Spy." The former differ in certain notable particulars from the human teeth

of to-day, whilst the latter have already become almost identical with those of some primitive races now existing. The author considers that we are now in process of losing our third molars and lateral incisors.

(170) IMMIGRATION COMMISSION. Changes in Bodily Form of Descendants of Immigrants. Washington, 1910.

The peoples considered in this report are the East European Hebrews, Sicilians and Calabrians in their own country and again in the districts of New York City. It is concluded that the cephalic index differs markedly in the European born and in the American born Hebrews and Sicilians, and that the descendants of these two quite distinct races when born in America are very much alike in respect to this character. Also, whilst the physique of the Hebrew improves in the American environment that of the Sicilian deteriorates. The treatment is rather inconclusive and somewhat inadequate.

(171) JACKMANN, O. Der Einfluss der Mikroben auf die Entstehung der Menschenrassen. Arch. f. Rassen- und Gesells. Biol., Bd. VI, S. 754—797. Leipzig, 1909.

(172) JACOB, S. M. On the Correlation of Areas of Matured Crops and the Rainfall. Mem. Asiatic Soc. of Bengal, Vol. II, pp. 347—429. Calcutta, 1910.

(173) ——. Inbreeding in a Stable Simple Mendelian Population with Special Reference to Cousin Marriage. Proc. Roy. Soc., B, Vol. LXXXIV, pp. 23—42. London, 1911.

(174) JAKOB, C. and ONELLI, C. Vom Tierhirn zum Menschenhirn. Vergleichend morphologische, histologische und biologische Studien zur Entwickelung der Grosshirn-Hemisphären und ihrer Rinde. J. F. Lehmann, Munich, 1911.

(175) JENNINGS, H. S. Experimental evidence on the effectiveness of Selection. Amer. Naturalist, Vol. XLIV, pp. 136—145. New York, 1910.

(176) JENNINGS, H. S. and HARGITT, G. T. Characteristics of the diverse Races of Paramaecium. Jour. of Morph., Vol. XXI, pp. 496—561. Philadelphia, 1910.

(177) JENNINGS, H. S. Pure lines in the Study of Genetics in lower organisms. Amer. Naturalist, Vol. XLV, pp. 79—89. New York, 1911.

(178) JOHANNSEN, W. Elemente der exakten Erblichkeitslehre. G. Fischer, Jena, 1909.

(179) JOHNSON, MYRTLE E. A quantitative study of the development of the chain in Salpa fusiformis-runcinata. Univ. of California Publications, Zoology, Vol. VI, pp. 145—176. Berkeley, 1910.

(180) JOHNSON, R. H. Determinate Evolution in the Color Patterns of the Lady Beetles. Carnegie Inst., Washington, Pub. No. 122. 1910.

(181) JORDAN, H. E. Recent Literature touching the question of Sex Determination. Amer. Naturalist, Vol. XLIV, pp. 245—252. 1910.

(182) KAMMERER, P. Direkt induzierte Farbanpassung und deren Vererbung. Zeitschr. f. indukt. Abstam.- u. Vererbungslehre, Bd. IV, S. 279—288. Berlin, 1911.

(183) ——. Vererbung künstlicher Farbenveränderungen. Die Umschau, Dec. 1909.

(184) KING, HELEN D. The Effects of Various Fixatives on the Brain of the Albino Rat, with an Account of a Method of Preparing this Material for a Study of the Cells in the Cortex. Anat. Record, Vol. IV, pp. 213—244. Philadelphia, 1910.

(185) ——. Temperature as a Factor in the Determination of Sex in Amphibians. Biol. Bull., Vol. XVIII, pp. 131—137. Philadelphia, 1910.

(186) ——. Studies on Sex-determination in Amphibians. IV. The Effects of External Factors, acting before or during the time of Fertilization, on the Sex Ratio of Bufo lentiginosus. Biol. Bull., Vol. XX, pp. 205—235. Philadelphia, 1911.

(187) KING, HELEN D. The Effects of Pneumonia and of Post-mortem Changes on the Percentage of Water in the Brain of the Albino Rat. Jour. of Comp. Neurol. and Psych., Vol. XXI, pp. 147—154. Philadelphia, 1911.

(188) ——. The Effects of Semi-Spaying and of Semi-Castration on the Sex Ratio of the Albino Rat (*Mus Norvegicus Albinus*). Jour. of Exper. Zoology, Vol. X, pp. 381—392. Baltimore, 1911.

(189) KLEBS, G. Alterations in the Development and Forms of Plants as a Result of Environment. Croonian Lecture, Proc. Roy. Soc., B, Vol. LXXXII. London, 1910.

(190) KNIBBS, G. H. Studies in Statistical Representation. On the Nature of the Curve $y = Ax^m e^{nx^p}$. Jour. Roy. Soc. of New South Wales, Vol. XLIV, pp. 341—367. Sydney, 1911.

(191) ——. Studies in Statistical Representation. Statistical Applications of the Fourier Series. Illustrated by the analysis of the rates of marriage, temperature, suicide, &c. Jour. Roy. Soc. of New South Wales, Vol. XLV, pp. 76—110. Sydney, 1911.

(192) KOLLMANN, J. Das Problem der Gleichheit der Rassen. Arch. f. Rassen- u. Gesells. Biol., Bd. VIII, S. 339—348. Leipzig, 1911.

(193) KORIBA, K. Ueber die individuelle Verschiedenheit in der Entwickelung einiger fortwachsenden Pflanzen mit besonderer Rücksicht auf die Aussenbedingungen. Jour. of Coll. of Sci. Imp. Univ. of Tokyo, Vol. XXVII, Art. 3. Tokyo, 1909.

(194) KÖRÖSY, K. Studien über Puls- und Atmungsfrequenz. Deutsch. Arch. f. klinische Medizin, Bd. CI, S. 267—282. Leipzig, 1910.

(195) LAMB, G. On the spread of Epidemic Plague through Districts with Scattered Villages. Jour. of Hygiene, Plague Number, Vol. X, pp. 350—443. Cambridge, 1910.

(196) LEBAS, R. Études critiques des Stigmates Anatomiques de la Criminalité et de quelques Théories Criminalistes Actuelles. Jouve & Co., Paris, 1910.

(197) LEHMANN, E. Über Zwischenrassen in der Veronica-Gruppe Agrestis. Zeitschr. f. induk. Abstam.- und Vererbungslehre, Bd. II, S. 145—208. Berlin, 1909.

(198) ——. Über Merkmalseinheiten in der Veronika-Sektion Alsinebe. Zeitsch. f. Botanik, Bd. II, S. 577—602. Jena, 1910.

(199) LLOYD, R. E. A Description of the Deep-Sea Fish caught by the R.I.M.S. Ship *Investigator* since the year 1900, with supposed evidence of mutation in Malthopsis. Memoirs of the Indian Museum, Vol. II, pp. 139—180. Calcutta, 1909.

(200) ——. The Races of Indian Rats. Records of Indian Museum, Vol. III, pp. 1—100. Calcutta, 1909. Vol. V, pp. 105—113. 1910.

(201) ——. The Relation between fertility and normality in Rats. Records of Indian Museum, Vol. III, pp. 261—265. Calcutta, 1909.

(202) LOBSIEN, M. Korrelationen zwischen den unterrichtlichen Leistungen einer Schülergruppe. Zeitschr. f. exp. Pädagog., Bd. XI, S. 146—164. Leipzig, 1911.

(203) LOTH, E. Über die Neuerungen in der Diagraphentechnik. Korres. Blatt der Deutsch. Gesell. f. Anthrop., Ethnol. und Urgesch., Brunswick, Bd. XL, S. 26—28. 1909.

(204) ——. Anthropologische Untersuchungen über das Hautleistensystem der Polen. Zeitschr. f. Morph. u. Anthrop., Bd. XIII, S. 77—96. Stuttgart, 1910.

(205) ——. Beiträge zur Kraniologie der Polen. Zeit. f. Morph. u. Anthrop., Bd. XIV, S. 305—338. Stuttgart, 1911.

(206) LOTKA, A. J. Evolution in Discontinuous Systems. Jour. of Washington Academy of Sciences, Vol. II, pp. 2—6, 49—59 and 66—74. Washington, 1912.

(207) LOVE, H. H. Are Fluctuations Inherited? Amer. Naturalist, Vol. XLIV, pp. 412—423. New York, 1910.

(208) Lovejoy, A. O. The Argument for Organic Evolution before "The Origin of Species." Pop. Sci. Mon., pp. 499—550. New York, 1909.

(209) Lutz, A. and Neiva, A. Erephopsis Auricincta. Eine neue Tabinidenart aus der Subfamilie Pangoninae. Mem. Inst. Oswaldo Cruz, Rio de Janeiro, Vol. i, pp. 12—13. 1909.

(210) Lutz, F. E. Experiments with Drosophila ampelophila concerning Evolution. Carnegie Inst., Washington, Pub. No. 143. 1911.

(211) Macdougal, D. T. Organic Response. Amer. Naturalist, Vol. xlv, pp. 1—40. New York, 1911.

(212) March, L. Pour la Race Infertilité et Puériculture. Rev. d'Hygiène et de Méd. Infantiles, T. ix, Nos. 5, 6. Paris, 1910.

(213) ——. Essai sur une mode d'exposer les principaux éléments de la théorie statistique. Jour. de la Soc. de Stat. de Paris. Année li, pp. 417—486. Nancy, 1910.

(214) Martius, F. Pathogenese innerer Krankheiten. F. Deuticke, Leipzig, 1909.
(A good résumé forcibly expressed.)

(215) McClendon, J. F. On the Effect of External Conditions on the Reproduction of Daphnia. Amer. Naturalist, Vol. xliv, pp. 404—411. New York, 1910.

(216) Meijere, J. C. H. de. Über getrennte Vererbung der Geschlechter. Arch. f. Rassen- u. Gesells. Biol. 1911, Bd. viii, S. 553—603, 697—752. Leipzig, 1911.

(217) Meldola, R. Evolution—Darwinian and Spencerian. The Herbert Spencer Lecture, Oxford, 1910.

(218) Merrins, E. M. Anthropometry of Chinese Students. Chinese Med. Jour., Sept. 1910.

A discussion of certain physical measurements taken on 219 Chinese boys and 69 Chinese girls in the Wuchang Schools and also on Chinese boys of the Shantung Union College. The author concludes that (1) There is a great need for proper registration and publication of vital statistics in China, (2) The physical development of Chinese children is behind that of American or English children, (3) There should be elimination of the morally weak from Chinese Schools. Dr Merrins also finds from an investigation of 200 Chinese families a striking preponderance of male births over female—of 1230 births there were 788 males and only 442 females.

(219) Minot, C. S. The Method of Science. Science, N. S., Vol. xxxiii, pp. 119—131. New York, 1911.

(220) Moenkhaus, W. J. The effects of inbreeding and selection on the fertility, vigor and sex-ratio of Drosophila ampelophila. Jour. of Morph., Vol. xxii, pp. 123—154. Philadelphia, 1911.

(221) Mollison, T. Die Körperproportionen der Primaten. Morph. Jahrb., Bd. xlii, S. 79—304. Leipzig, 1910.

(222) Montgomery, T. The influence of changed external conditions on the development of moths. Amer. Naturalist, Vol. xlv, pp. 364—371. New York, 1911.

(223) Morgan, T. H. A biological and cytological study of sex determination in phylloxerans and aphids. Jour. Exp. Zool., Vol. vii, pp. 239—352. Baltimore, 1909.

(224) ——. Chance or purpose in the origin and evolution of adaptation. Science, N. S., Vol. xxxi, pp. 201—210. New York, 1910.

(225) ——. Sex Limited Inheritance in Drosophila. Science, N. S., Vol. xxxii, pp. 120—122. New York, 1910.

(226) ——. Chromosomes and Heredity. Amer. Naturalist, Vol. xliv, pp. 449—496. New York, 1910.

(227) Morgan, T. H. Experiments bearing on the Nature of the Karyokinetic Figure. Proc. Soc. for Exp. Biol. and Med., Vol. vii, No. 4, p. 132. New York, 1910.

(228) ——. The application of the conception of pure lines to sex-limited inheritance and to sexual dimorphism. Amer. Naturalist, Vol. xlv, pp. 65—78. New York, 1911.

(229) ——. An attempt to analyse the Constitution of the Chromosomes on the basis of sex-limited inheritance in Drosophila. Jour. Exp. Zool., Vol. xi, pp. 365—413. New York, 1911.

(230) Mortara, G. Tavola di sopravvivenza e delle variazioni di stato civile e tavola di natalità legittima per la popolazione femminile del Comune di Milano, 1899—1902. Atti del R. Inst. d'Incoraggiamento di Napoli, Ser. vi, T. v, pp. 253—281. Naples, 1908.

(231) ——. Nozioni elementari intorno ad alcune categorie di rapporti statistici. Giorn. d. Econ. e Riv. di Stat., Rome, T. xl, Ser. iii, pp. 217—234. Naples, 1910.

(232) ——. Metodi elementari per lo Studio delle distribuzioni di caratteri. Giorn. d. Econ. e Riv. di Stat., Rome, T. xli, Ser. iii, pp. 9—25. 1910.

(233) ——. I concepimenti antenuzziali. Giorn. degli Econ. e Rivista di Statistica, T. xliii, Ser. iii, pp. 105—204. Rome, 1911.

(234) ——. L'incubo dello spopolamento e l'Italia. D'Angelo, Messina, 1912.

The Author adds a note on the decrease in the birth-rate in Italy, giving 23 tables of statistics.

(235) Mott, F. W. Alcohol and Insanity. Brit. Jour. of Inebriety, Vol. ix, pp. 5—17. London, 1911.

(236) ——. Tuberculosis in the London County Asylums. Archives of Neurology and Psychiatry, Vol. iv, pp. 71—115. London, 1909.

(237) ——. The Inborn Factors of Nervous and Mental Disease. Brain, Vol. xxxiv, pp. 73—101. London, 1911.

(238) ——. Heredity and Insanity. Eugenics Review, pp. 1—25. London, 1911.

Contains 19 pedigrees of mentally defective stocks. The Author concludes that hereditary predisposition is the most important factor in the production of insanity.

(239) Nettleship, E. On some Hereditary Diseases of the Eye. The Bowman Lecture. Trans. Ophthal. Soc., Vol. xxix, pp. lvii—cxcviii. London, 1909.

(240) ——. Seven new pedigrees of hereditary cataract. Trans. Ophthal. Soc., Vol. xxix, pp. 188—211. London, 1909.

(241) ——. Some Points in Relation to the Heredity of Disease. St Thomas's Hosp. Gaz., Vol. xx, pp. 37—65. London, 1910.

(242) Newsholme, A. Report on Infant and Child Mortality. Ann. Rep. Local Gov. Board, London, 1909-10.

(243) Nilsson-Ehle, H. Kreuzungsuntersuchungen an Hafer und Weizen. Lund, 1909.

(244) Osborn, H. F. The Continuous Origin of Certain Unit Characters as Observed by a Palaeontologist. Amer. Naturalist, Vol. xlvi, pp. 185—206, 249—278. New York, 1912.

(245) Papillault, G. Sur quelques erreurs de méthode en Criminologie. Rev. de l'École d'Anthrop. de Paris, T. xx, pp. 321—334. Paris, 1910.

(246) ——. Galton et la Bio-Sociologie. Rev. Anthropologique, Paris, T. xxi, pp. 56—65. 1911.

(247) Pearl, R. Studies on the Physiology of Reproduction in the Domestic Fowl. i. Regulation in the Morphogenetic Activity of the Oviduct. Jour. of Exp. Zool., Vol. vi, pp. 339—359. Baltimore, 1909.

(248) PEARL, R. Inheritance of Hatching Quality of Eggs in Poultry. Amer. Breeder's Mag., Vol. I, pp. 129—133. Washington, 1910.

(249) ——. Breeding for Production in Dairy Cattle in the Light of Recent Advances in the Study of Inheritance. 8th Annual Rep. of Commissioners of Agriculture of the State of Maine, pp. 118—129. Orono, Maine, 1910.

(250) ——. Breeding Poultry for Egg Production. Ann. Rep. Maine Agr. Exp. Stat., pp. 113—176. Orono, Maine, 1911.

(251) ——. Data on the Relative Conspicuousness of Barred and Self-Colored Fowls. Amer. Naturalist, Vol. XLV, pp. 107—117. New York, 1911.

(252) ——. Inheritance of Fecundity in the Domestic Fowl. Amer. Naturalist, Vol. XLV, pp. 321—345. New York, 1911.

(253) ——. Some recent Studies on Variation and Correlation in Agricultural Plants. Amer. Naturalist, Vol. XLV, pp. 415—425. New York, 1911.

(254) ——. The Personal Equation in Breeding Experiments involving certain Characters of Maize. Biol. Bull., Vol. XXI, pp. 339—366. New York, 1911.

(255) ——. Inheritance in "Blood Lines" in breeding animals for Performance, with special Reference to the "200-egg hen." Ann. Rep. Amer. Breeder's Assoc., Vol. VI, pp. 317—326. Washington, 1911.

(256) PEARL, R. and BARTLETT, J. M. The Mendelian Inheritance of certain Chemical Characters in Maize. Zeitschr. für induk. Abstam.- und Vererbungslehre, Bd. VI, S. 4—28. Berlin, 1911.

(257) PEARL, R. and CURTIS, M. R. Studies on the Physiology of Reproduction in the Domestic Fowl. III. A Case of Incomplete Hermaphroditism. Biol. Bull., Vol. XVII, pp. 271—284. Philadelphia, 1909.

(258) PEARL, R. and SURFACE, F. M. Selection Index Numbers and their use in Breeding. Amer. Naturalist, Vol. XLIII, pp. 385—400. New York, 1909.

(259) ——. Poultry Notes. Maine Agr. Exp. Stat. Bull., No. 165, pp. 29—48. Orono, Maine, 1909.

(260) ——. Data on the Inheritance of Fecundity obtained from the records of Egg Production of the Daughters of "200-egg" Hens. Maine Agr. Exp. Stat. Bull., No. 166, pp. 49—84. Orono, Maine, 1909.

(261) ——. Data on certain Factors influencing the Fertility and Hatching of Eggs. Maine Agr. Exp. Stat. Bull., No. 168, pp. 105—164. Orono, Maine, 1909.

(262) ——. Is there a Cumulative effect of Selection? Zeitschr. f. induk. Abstamm. u. Vererbungs., Vol. II, pp. 257—275. Berlin, 1909.

(263) ——. Studies on Hybrid Poultry. Ann. Rep. Maine Agr. Exp. Stat., pp. 84—116. Orono, Maine, 1910.

(264) ——. Experiments in Breeding Sweet Corn. Ann. Rep. Maine Agr. Exp. Stat., pp. 249—307. Orono, Maine, 1910.

(265) ——. On the Inheritance of the Barred Color Pattern in Poultry. Arch. Entwicklungsmech. Leipzig, Bd. XXX, S. 45—61. Leipzig, 1910.

(266) ——. Further data regarding the Sex-limited Inheritance of the Barred Color Pattern in Poultry. Science, N. S., Vol. XXXII, pp. 870—874. New York, 1910.

(267) ——. A Biometrical Study of Egg Production in the Domestic Fowl. I. Variation in Annual Egg Production. II. Seasonal Distribution of Egg Production. U. S. Dep. of Agr. Bureau of Animal Industry, Bull. 110, pp. 1—80. 1909. Bull. 110, Part II, pp. 81—170. Washington, 1911.

(268) Pearson, K., Nettleship, E. and Usher, C. H. Albinism in Man. Part i. Biometric Laboratory Publications, Biometric Series, vi. Dulau & Co., Ltd. London, 1911.

(269) Pearson, K. The Fight against Tuberculosis and the Death-rate from Phthisis. Biometric Laboratory Publications, Questions of the Day and of the Fray, iv. Dulau & Co., Ltd. London, 1911.

(270) ——. On a novel method of regarding the Association of Two Variates classed solely in Alternate Categories. Biometric Laboratory Publications, Biometric Series, viii. Dulau & Co., Ltd. London, 1912.

(271) ——. Social Problems: Their Treatment, Past, Present, and Future. Biometric Laboratory Publications, Questions of the Day and of the Fray, No. v. Dulau & Co., Ltd. London, 1912.

(272) ——. Tuberculosis, Heredity and Environment. Galton Eugenics Laboratory Lectures, viii. Dulau & Co., Ltd. London, 1912.

(273) Phillips, J. C. Size inheritance in ducks. Jour. Exp. Zool., Vol. xii, pp. 369—381. Baltimore, 1912.

(274) Plate, L. Vererbungslehre und Deszendenztheorie. Festschrift für Hertwig, Bd. ii, S. 538—610. Jena, 1910.

(275) ——. Ein Versuch zur Erklärung der gynephoren Vererbung menschlicher Erkrankungen. Arch. f. Rassen- u. Gesells. Biol., Bd. viii, S. 164—171. Leipzig, 1911.

(276) Ploetz, A. Ziele und Aufgaben der Rassenhygiene. Deutsches Vierteljahrsschr. f. öffentl. Gesundh., Bd. xliii, S. 164—199. Brunswick, 1910.

(277) Poniatowski, S. Über den Wert der Index-Klassifikationen. Archiv f. Anthrop., Bd. x, S. 1—7. Brunswick, 1911.

(278) Ramaley, F. Mendelian Proportions and the Increase of Recessives. Amer. Naturalist, Vol. xlvi, pp. 344—351. New York, 1912.

(279) Renvall, A. Die periodischen Erscheinungen der Reproduktion der Kiefer an der polaren Waldgrenze. Helsingfors, 1912.

(280) Riddle, O. Studies with Sudan III. in Metabolism and Heredity. Jour. of Exp. Zool., Vol. viii, pp. 163—184. Baltimore, 1910.

(281) Rignano, E. Le Rôle des "Théoriciens" dans les Sciences Biologiques et Sociologiques. Scientia, Vol. xi, pp. 218—231. Bologna, 1912.

(282) Ritter, W. E. and Johnson, M. E. The Growth and Differentiation of the Chain of Cyclosalpa Affinis Chamisso. Jour. of Morph., Vol. xxii, pp. 396—444. Philadelphia, 1911.

(283) Rivers, W. H. R. The Genealogical Method of Anthropological Inquiry. Sociological Review, Vol. iii, pp. 1—12. Manchester, 1910.

(284) Robertson, A. W. D. Craniological Observations on the Lengths, Breadths and Heights of a Hundred Australian Aboriginal Crania. Proc. of Roy. Soc. of Edin., Vol. xxxi, pp. 1—16. Edinburgh, 1910–11.

(285) Röse, C. Über die mittlere Durchbruchszeit der bleibenden Zähne des Menschen. Deutsche Monatsschrift für Zahnheilkunde, Jahrg. xxvii, S. 553—570. Berlin, 1909.

(286) Ross, E. H. The Influence of certain biological Factors on the question of the Migration of Mosquitoes. Jour. of Trop. Med. and Hygiene, Vol. xii, pp. 256—258. London, 1909.

(287) Russo, A. Studien über die Bestimmung des weiblichen Geschlechtes. Jena, 1909.

(288) Sasse, J. Wie sollen wir urteilen über die Grösse der drei Hauptdurchmesser am menschlichen Schädel? Zeit. f. Morph. u. Anthrop., Bd. xii, pp. 559—574. Stuttgart, 1910.

(289) Schallmayer, W. Generative Ethik. i. Das Verhältnis des generativen Ethik zur Ethik überhaupt und zur sozialen Sexualethik. ii. Die Notwendigkeit eines generativen Ethik und der Weg zu ihrer Entwicklung. Archiv f. Rassen- und Gesells. Biol., Bd. vi, S. 199—231. Leipzig, 1909.

(290) Schwalbe, G. Ueber Darwins Werk: "Die Abstammung der Menschen." Zeitschr. f. Morph. u. Anthrop., Bd. xii, S. 441—472. Stuttgart, 1910.

(291) Sheppard, W. F. The Accuracy of Interpolation by Finite Differences. Proc. Lond. Math. Soc., Ser. ii, Vol. x, pp. 139—172. London, 1911.

(292) ——. Summation of the Coefficients of some Terminating Hypergeometric Series. Proc. Lond. Math. Soc., Ser. ii, Vol. x, pp. 469—478. London, 1911.

(293) Shull, G. H. A Simple Chemical Device to Illustrate Mendelian Inheritance. The Plant World, Vol. xii, 1909, pp. 145—153.

(294) ——. The "presence and absence" Hypothesis. Amer. Naturalist, Vol. xliii, pp. 410—419. New York, 1909.

(295) Shull, A. F. Studies in the Life Cycle of Hydatina Senta. i. Artificial Control of the Transition from the Parthenogenetic to the Sexual Method of Reproduction. Jour. of Exper. Zool., Vol. viii, pp. 311—354. Baltimore, 1910.

(296) Shull, G. H. Inheritance of Sex in Lychnis. Bot. Gaz., Vol. xlix, pp. 110—125. Chicago, 1910.

(297) ——. Hybridization Methods in Corn Breeding. Amer. Breeder's Mag., Vol. i, pp. 98—107. Washington, 1910.

(298) ——. Germinal Analysis through Hybridization. Proc. Amer. Phil. Soc., Vol. xlix, pp. 281—290. Philadelphia, 1910.

(299) ——. Color Inheritance in Lychnis Dioica L. Amer. Naturalist, Vol. xliv, pp. 83—91. New York, 1910.

(300) ——. Reversible Sex-mutants in Lychnis Dioica. Bot. Gaz., Vol. lii, pp. 329—368. Chicago, 1911.

(301) ——. The Genotypes of Maize. Amer. Naturalist, Vol. xlv, pp. 234—252. New York, 1911.

(302) Siebert, F. Die Rassenidee und die liberale Weltanschauung. Arch. f. Rassen- u. Gesells. Biol., Bd. viii, S. 198—223. Leipzig, 1911.

(303) Smith, G. Elliot. The People of Egypt. The Cairo Scientific Journal, pp. 51—63. Cairo, 1909.

Somewhat dogmatic treatment with no statistical data properly reduced.

(304) Smith, G. Elliot and Derry, D. E. The Archaeological Survey of Nubia. Bull. No. 6. Anatomical Report, dealing with the work during the months of Jan. and Feb. 1910. Cairo, 1910.

No proper reduction of statistical data.

(305) Smith, G. M. C. Heredity as a Causative Factor in Disease. Edin. Med. Jour., N. S., Vol. vi, pp. 422—429. Edinburgh, 1911.

(306) Snow, E. C. On the Intensity of Natural Selection in Man. Biometric Laboratory Publications, Studies in National Deterioration, vii. Dulau & Co., Ltd. London, 1911.

(307) ——. The Application of the Method of Multiple Correlation to the Estimation of Post-Censal Populations. Jour. of Roy. Stat. Soc., Vol. lxxiv, pp. 575—629. London, 1911.

The Author claims for the method of multiple correlation that it gives estimates of post-censal populations agreeing more closely with the true values than those obtained by any other method hitherto employed.

(308) Spearman, C. Correlation Calculated from Faulty Data. Brit. Jour. of Psych., Vol. iii, pp. 271—295. Cambridge, 1910.

Several highly disputable statements are made.

(309) Spillman, W. J. Mendelian phenomena without De Vriesian theory. Amer. Naturalist, Vol. xliv, pp. 214—228. New York, 1910.

(310) ——. A Theory of Mendelian Phenomena. Amer. Breeder's Mag., pp. 113—125. Washington, 1910.

(311) Stainer, E. The Hereditary Transmission of Defects in Man. Oxford, 1910.

(312) Stansfield, T. E. K. Heredity and Insanity. Jour. of Mental Science, Vol. lxvii, pp. 55—63. London, 1911.

(313) Steinmetz, S. R. L'Avvenire della razza. Rivista Italiana di Sociologia, Anno xiv, pp. 485—509, fasc. v, vi. Rome, 1910.

(314) Stockard, C. R. The Influence of Alcohol and other Anaesthetics on Embryonic Development. Amer. Jour. of Anat., Vol. x, pp. 369—392. Philadelphia, 1910.

(315) Stolyhwo, K. Der Osteophor-Projektiometer. Korres.-Blatt der Deutsch. Gesell. f. Anthrop., Ethnol. u. Urgesch., Bd. xli, S. 25—30. Brunswick, 1910.

(316) ——. W Sprawie kształtów goryloidycznych i orangoidycznych. Warsaw Society of Sciences Proc., pp. 69—79. Warsaw, 1911.

(317) ——. Zur Frage einer neuen polygenistischen Theorie der Abstammung des Menschen. Zeitschr. f. Ethnol., Bd. xliv, S. 97—104. Berlin, 1912.

(318) Strasburger, E. Ueber geschlechtbestimmende Ursachen. Jahrb. f. wiss. Bot., Bd. xlviii, S. 427—520. Leipzig, 1910.

(319) Sturrock, J. P. The Modern Aspects of Eugenics. The Caledonian Med. Jour., Vol. viii, pp. 357—369. Glasgow, 1911.

(320) Sturtevant, A. H. On the Inheritance of Color in the American Harness Horse. Biol. Bull., Vol. xix, pp. 204—216. Philadelphia, 1910.

(321) Sumner, F. B. Some effects of external conditions upon the white mouse. Jour. of Exp. Zool., Vol. vii, pp. 97—155. Baltimore, 1909.

(322) ——. The Reappearance in the offspring of artificially produced Parental Modifications. Amer. Naturalist, Vol. xliv, pp. 1—18. New York, 1910.

(323) ——. An Experimental Study of Somatic Modifications and their Reappearance in the Offspring. Arch. f. Entwicklungsmech., Bd. xxx, S. 317—348. Leipzig, 1910.

(324) Tammes, T. Das Verhalten fluktuierend variierender Merkmale bei der Bastardierung. Recueil des Travaux botaniques Néerlandais, T. viii, pp. 201—284. Nimeguen, 1911.

(325) Tennent, D. H. The Dominance of Maternal or of Paternal Characters in Echinoderm Hybrids. Arch. f. Entwicklungsmech., Bd. xxix, S. 1—14. Leipzig, 1910.

(326) Thiele, T. N. Interpolationsrechnung. B. C. Teubner, Leipzig. 1909.

(327) Thomson, D. A research into the Cranial Measurements of the Insane, comparing them with those of the Sane. Journal of Mental Science, Vol. lv, pp. 280—282. London, 1909.

(328) Tocher, J. F. The Necessity for a National Eugenic Survey. Eugenics Review, pp. 124—141. London, 1910.

(329) Tower, W. L. The Determination of Dominance and the Modification of Behavior in Alternative (Mendelian) Inheritance, by Conditions Surrounding or Incident upon the Germ Cells of Fertilization. Biol. Bull., Vol. xviii, pp. 285—352. Philadelphia, 1910.

(330) TROUP, McD. and MAYNARD, G. D. Modern Statistical Methods. The Lancet, Vol. I, pp. 1336—1343. London, 1910.

(331) TURNER, G. A. Some of the Tribal Marks of the South African Native Races. Transvaal Med. Jour., Vol. VI, pp. 141—153. Johannesburg, 1911.

(332) VOGLER, P. Probleme und Resultate variationsstatistischer Untersuchungen an Blüten und Blütenständen. Jahrb. der St Gallischen naturwiss. Gesells. St Gallen, 1910. (Offprint.)

(333) ——. Variation der Anzahl der Strahlblüten bei einigen Kompositen. Bot. Centralb., Bd. XXV, S. 387—396. Dresden, 1910.

(334) ——. Die Variation der Blattspreite bei Cytisus laburnum L. Bot. Centralb., Bd. XXVII, S. 391—437. Dresden, 1911.

(335) ——. Neue variationsstatistische Untersuchungen an Compositen. Jahrb. d. naturwiss. Gesells., pp. 1—32. St Gallen, 1911.

(336) WALTER, H. E. Variations in Urosalpinx. Amer. Naturalist, Vol. XLIV, pp. 576—594. New York, 1910.

(337) WARREN, E. Notes on the Life Histories of Natal Termites, Based on the Observations of the late Mr G. D. Haviland. Annals of Natal Gov. Mus., Vol. II, pp. 113—128. London, 1909.

(338) WEBB, A. D. Notes on some difficulties met with in international statistical comparisons. Jour. Roy. Stat. Soc., Vol. LXXIII, pp. 10—25. London, 1910.

(339) WEINBERG, W. Die Anlage zur Mehrlingsgeburt beim Menschen und ihre Vererbung. Arch. f. Rassen- und Gesells. Biol., Bd. VI, S. 322—339, 470—482, 610—629. Leipzig, 1909.

(340) ——. Zur Statistik der Tuberkulose beider Ehegatten. Zeit. f. soz. med. Säuglingsfürsorge und Krankenhauswesen, Bd. V, S. 309—316. Leipzig, 1910.

(341) ——. Weitere Beiträge zur Theorie der Vererbung. Arch. f. Rassen- u. Gesells. Biol., Bd. VII, S. 35—49. Leipzig, 1910.

(342) ——. Die rassenhygienische Bedeutung der Fruchtbarkeit. Arch. f. Rassen- u. Gesells. Biol., Bd. VIII, S. 25—32. Leipzig, 1911.

(343) WHITNEY, D. D. The Influence of External Conditions upon the Life Cycle of Hydatina Senta. Science, Vol. XXXII, pp. 345—349. New York, 1910.

(344) ——. The Effects of Alcohol not Inherited in Hydatina Senta. Amer. Naturalist, Vol. XLVI, pp. 41—56. New York, 1912.

(345) ——. Reinvigoration Produced by Cross-Fertilization in Hydatina Senta. Jour. Exp. Zool., Vol. XII, pp. 337—362. Baltimore, 1912.

(346) ——. "Strains" in Hydatina Senta. Biol. Bull., Vol. XXII, pp. 205—217. Philadelphia, 1912.

(347) WILSON, E. B. The chromosomes in relation to the determination of sex. Science Prog., Vol. V, pp. 570—592. London, 1910.

(348) WILSON, J. The Colours of Highland Cattle. Sci. Proc. Roy. Dublin Soc., Vol. XII, pp. 66—76. Dublin, 1909.

(349) ——. The Separate Inheritance of Quantity and Quality in Cows' Milk. Sci. Proc. Roy. Dublin Soc., Vol. XII, pp. 470—479. Dublin, 1910.

(350) WINCH, W. H. The Transfer of Improvement in Memory in School Children. Brit. Jour. of Psych., Vol. III, pp. 386—405. Cambridge, 1910.

(351) WINKLER, H. Ueber die Nachkommenschaft der Solanum. Propfbastarde und die Chromosomenzahlen ihrer Keimzellen. Zeitschr. f. Bot., Bd. I, S. 1—38. Jena, 1909.

(352) WINKLER, H. Weitere Untersuchungen über Propfbastarde. Zeitschr. f. Bot., Bd. I, S. 315—345. Jena, 1909.

(353) WINSLOW, C. E. A. A Statistical Criterion for Species and Genera among the Bacteria. Bull. of Torrey Botanical Club, Vol. XXXVI, pp. 31—39. Boston, U.S.A., 1909.

(354) WINSLOW, C. E. A. and WALKER, L. T. A Case of Non-inheritance of Fluctuating Variations among Bacteria. Jour. of Infectious Diseases, Vol. VI, pp. 90—97. Chicago, 1909.

(355) WOODS, F. A. The Birthplaces of Leading Americans and the Question of Heredity. Science, N. S., Vol. XXX, pp. 17—21. New York, 1909.

(356) ——. American Men of Science and the Question of Heredity. Science, N. S., Vol. XXX, pp. 205—210. New York, 1909.

(357) ——. Laws of Diminishing Environmental Influences. Pop. Sci. Mon., Vol. LXXVII, pp. 313—336. New York, 1910.

(358) ——. Historiometry as an Exact Science. Science, N. S., Vol. XXXIII, pp. 568—574. New York, 1911.

(359) YULE, G. U. On the distribution of deaths with age when the causes of death act cumulatively, and similar frequency distributions. Jour. Roy. Stat. Soc., Vol. LXXIII, pp. 26—38. London, 1910.

(360) ——. On the Interpretation of Correlations between Indices or Ratios. Jour. Roy. Stat. Soc., Vol. LXXIII, pp. 644—647. London, 1910.

The author makes a very curious blunder in criticising the theory of spurious correlation.

(361) ——. An Introduction to the Study of Statistics. Griffin & Co., London. 1911.

(362) ZALĘSKI, E. I Zastosowanie Wielokatów Częstotliwości do selekcyi roślin. Roczniki Nauk Rolniczych, Vol. IV, pp. 243—274. Cracow, 1909.

On the application of frequency-polygons to plant selection with somewhat sweeping statements as to the interpretation of asymmetry.

ON THEORIES OF ASSOCIATION.

By KARL PEARSON, F.R.S. and DAVID HERON, D.Sc.

CONTENTS.

(1) *Introductory.*

THE recent paper by Mr Yule entitled "On the Methods of Measuring Association between Two Attributes*" calls for an early reply on two grounds,—first because its singularly acrimonious tone is to us wholly inexplicable, not to say unusual, and secondly because we believe that, if Mr Yule's views are accepted, irreparable damage will be done to the growth of modern statistical theory.

* *Journal of the Royal Statistical Society*, Vol. LXXV, pp. 579—652. London, 1912.

Mr Yule has invented a series of statistical methods which are in no case based on a reasoned theory, but which possess the dangerous fascination of very easy and ready application, and therefore are at once seized upon as applicable to all sorts of problems by those who are without adequate training in statistical theory, or without the mathematical knowledge requisite to weigh cautiously their logical basis. The methods to which we refer are these:

(i) the use of the so-called "coefficient of association" to measure the relationship of two attributes;

(ii) the use of a new coefficient, which Mr Yule terms a "coefficient of colligation," apparently to be used in like cases with the coefficient of association;

(iii) the development of a method which first appeared in a paper by the late Mr John Gray*; in this method each group of a contingency table is considered as a cell of *unit* subrange for both variates. This assumption being made, Mr Yule calculates the coefficient of correlation by the product-moment method, and on the basis of this procedure terms his coefficient the coefficient of correlation and uses the customary letter r for it.

Such a terminology is absolutely unjustifiable and can only confuse the uninstructed and undiscerning reader. If the groups were extended from 5 or 8 to an indefinite number, all Mr Yule would reach by this method would be a correlation of *ranks*, not of variates. As it is, he has obtained a correlation of ranks with enormous "brackets†." It does not seem to have occurred to him that the correlation of ranks may be quite different from the true correlation of variates, and that in cases where we do know the relationship the correlation of ranks is sensibly lower than that of variates. Further, he makes no suggestion that a very fundamental correction—that of the variate and class-index correlation—is needful before this method could possibly be applied to deduce a limit to the true correlation of variates. For these reasons we shall term Mr Yule's latest method of approaching the problem of relationship of attributes the *method of pseudo-ranks*. We are concerned principally therefore in this paper with the so-called method of association and the method of pseudo-ranks. In addition we deal incidentally with other coefficients and reply to certain criticisms, not to say charges, Mr Yule has made against the work of one or both of us.

(2) *History of Subject.*

In view of the misstatements made in the discussion at the Royal Statistical Society, with regard to the history of the subject, a few preliminary remarks of an historical nature may be fitly made here. Mr Sanger, for example, said that "all statisticians before Mr Yule had this passion for the normal curve‡." This statement

* Gray and Tocher, *Journal of Anthrop. Institute*, 1900, Vol. xxx. p. 111.

† On the difficulty of "brackets" in the correlation of ranks, see Pearson, "On Further Methods of determining Correlation," *Drapers' Research Memoirs*, Dulau and Co., p. 36.

‡ *Journal Roy. Stat. Soc.* Vol. LXXV. p. 645.

is not only unfair to Perozzo, but to one of us, whom Mr Yule was directly attacking. Pearson's memoir on "Skew Variation in Homogeneous Material" was sent to the Royal Society on December 19, 1894. Mr Yule's memoir on Association was presented on October 20, 1899. He had previously attended the statistical lectures of Pearson and been an assistant to him during a period when nearly the whole work of the statistical laboratory turned on non-Gaussian distributions. A collection was then made of non-Gaussian material with a view to dealing with the correlation surfaces of continuous non-Gaussian variates. Among the material especially selected as providing extreme cases were (i) barometric heights (memoir presented 1897) and (ii) ages of husband and wife; of the latter the laboratory still possesses the contour lines drawn by Mr Yule under Pearson's direction to indicate extreme cases of what the latter has termed skew variation, or wide deviation from the normal surface. Another illustration of marked skewness is that of the contour lines of the correlation between the numbers of a particular suit in two partners' hands at whist. These latter curves were published by one of us in 1894, and at that time* it was distinctly stated that the contour curves for ages of husband and wife differed widely from the Gaussian type. It is singular that Mr Yule in the paper we are about to discuss should have made use of two of the extreme types of non-Gaussian frequency with which he was very familiar when he was an assistant in the University College Department of Applied Mathematics, and yet have allowed such a statement as that of Mr Sanger's to remain uncontradicted. The fact is that the promise made in 1894† to deal with skew correlation *surfaces* only remained unfulfilled because the differential equations to the surfaces obtained in that year have so far defied integration. Undoubtedly for continuous variates a generalised correlation surface should be the starting-point for attacking the problem of association‡. That it has remained unsolved shows only the extreme difficulty of the problem; it does not indicate that all statisticians before Mr Yule had "this passion for the normal curve."

And here we will at once emphasise the fundamental difference between Mr Yule and ourselves. Mr Yule, as we will indicate later, does not stop to discuss whether his attributes are really continuous or are discrete, or hide under discrete terminology true continuous variates. We see under such class-indices as 'death' or 'recovery,' 'employment' or 'non-employment' of mother, only measures of continuous variates—which of course are not *a priori* and necessarily Gaussian. Mr Greenwood in the discussion on Mr Yule's paper referred to the jibe§ about

* *Phil. Trans.* Vol. 186 A, p. 411.

† *Ibid.* p. 411.

‡ Such a surface, however, involves 5, 7, or more independent constants, not the *three* of the Gaussian surface, and as the fourfold table has, apart from its total, only three available constants, we could not hope to determine such a surface for a *fourfold* table without some additional knowledge or hypothesis, other than conveyed by the table itself, as to the nature of the frequency.

§ "We are considering," writes Mr Yule, "simply the performance as against the non-performance of the operation of vaccination. Similarly all those who have died of small-pox are equally dead: no one of them is more dead or less dead than another, and the dead are quite distinct from the survivors"

persons being 'dead' or 'not dead' and questioned whether Mr Yule was correct in treating the variate behind the class-index as discrete and not continuous. In the original paper of Pearson "strength to resist smallpox when incurred" was stated to be the variate, and all the evidence that has been produced since indicates its continuity; in precisely the same way vaccination and non-vaccination represent degrees of immunity in a continuous variate of which area of vaccination as indicated by extent of cicatrix and period since vaccination are contributory quantitative factors. Again "employment or non-employment of the mother" are not taken by us as signifying the presence or absence of a mere discrete attribute—for example whether she works in a factory or not—but as a class-index indicating that employed women, who have not only their home work but factory labour also, have on the whole more physical exertion to endure than those who are simply housewives. We are really seeking how far the continuous variate physical exertion of women affects infant welfare, and this is not a discrete variate any more than survival or death of infant is a discrete variate, when you view them merely as class-indices of physical fitness to survive in the child. In other words, for the great bulk of attributes, to which Mr Yule without analysis of their nature applies association, we should assert continuous variation. We hold therefore that in the main we are applying fourfold or other class divisions to continuous variates. Mr Yule thinks he has freed himself from all consideration of what the nature of this continuity may be; we consider his belief wholly fallacious. You cannot free yourself from some assumption as to the nature of the distribution when you are dealing with the association of attributes. And in ignorance of what the true distribution may be, what assumption will help you to the most probable result? On the basis of a very large experience of frequency curves and surfaces we have no hesitation in saying that up to the present time no distribution has been proposed which roundly represents experience so effectively as the Gaussian frequency. One of the present writers has indicated over and over again how it fails, and he has measured the significance of its failure, but has always recognised that he must put against this the large percentage of cases in which it gives reasonable results, close enough for all practical purposes*. Mr Yule

(*J. R. S. S.* Vol. LXXV. p. 612). Who the "we" are, Mr Yule does not tell us; but suppose "we" started to find the relation between age and place in an examination—say the mathematical tripos—should we learn more or less by treating the wranglers as a class-mark with no graduations and age 21 as a fixed division, or by assuming that the fourfold table: wrangler—not wrangler, minor—not minor, really covered continuous variations of age and class-place? Vaccination means vaccinated a week ago or forty years ago, a graduated immunity; the dead means a group who not only had no power to resist an attack of the given intensity, but in certain of the cases an attack of a far less intensity—it covers a class with graduated power of resistance. Mr Yule here as elsewhere is mastered by words, not seeking the realities behind classification.

* See, for example, C. D. Fawcett, *Biometrika*, Vol. I. p. 443; W. R. Macdonell, Vol. I. p. 184, Vol. III. p. 227; Pearson and Lee, Vol. II. pp. 362—7; R. Pearl, Vol. IV. p. 40; etc. etc. Compare, however, J. F. Tocher, Vol. V. p. 300, who for long series finds a certain amount of deviation from normality, generally in the direction of "leptokurtosis" (β_2 is large) not of much asymmetry (β_1 is small). Even in these cases we doubt whether any serious practical error would be introduced by the use of the Gaussian distribution, unless extreme dichotomies are made.

with a bias which may well be called in question has gone out of his way to select extreme cases, which had already been indicated by one of us as markedly non-Gaussian, but he makes no attempt to measure the wide range of physical characters for which the Gaussian is a legitimate practical assumption. Mr Yule refers to Dr Macdonell's memoir as a case in which the applicability of the Gaussian fourfold table method was "in the first place adequately tested" before adoption*. He leaves his uninitiated reader ignorant of two important facts, (i) that in the majority of fourfold classifications there is no possibility of such adequate testing because only the fourfold division has been provided, and (ii) the test in this case was directly made at the suggestion of Pearson and in his Laboratory to test the efficiency of the Gaussian method on ordinary data such as form probably nine-tenths of the frequencies which occur in practice. The work on this paper of Macdonell's began almost immediately on the completion of the theoretical memoir of 1900 on the fourfold table, and Mr Yule's statement that the warning of Pearson in the fundamental memoir of 1900 that normal correlation was not universal "seems to have been forgotten in a few weeks at most†" is, as many others of his statements from the historical standpoint, hopelessly inaccurate. Thus the paper on eye-colour in man and coat-colour in horses was presented in August, 1899, and *antedates* the presentation of the theoretical paper of February, 1900. The "warning" could hardly have been promptly forgotten, for the paper was withdrawn and rewritten *in order to test the value of the method of association then just propounded by Mr Yule, and to develop, when that was found defective, what, it was believed, was a better treatment.* Mr Yule writes that "Professor Pearson raised no objection then and as far as I know has raised no objection since to my coefficient Q; indeed he referred to 'the extreme elegance and simplicity of Mr Yule's coefficient of association.'" Naturally when one finds a method wholly inadequate one does not turn and rend an old pupil and former colleague. What Pearson did do was to test Mr Yule's Q against other similar coefficients and finding it less stable than any of them, it was dropped and has never been and never will be used in any work done under his supervision. But an interesting point arises here, which it is, perhaps, worth mentioning. Endeavouring to find for any fourfold division an analogue to Sheppard's median division formula, i.e. for a Gaussian fourfold

$$\begin{array}{c|c} a & b \\ \hline c & d \end{array}$$

$$r = \cos \pi \frac{b}{a+b} = \sin \frac{\pi}{2} \frac{a-b}{a+b},$$

Pearson hit upon the fact that the fourfold

$$\begin{array}{c|c} \sqrt{ad} & \sqrt{bc} \\ \hline \sqrt{bc} & \sqrt{ad} \end{array}$$

* *Journ. of R. S. S.* Vol. LXXV. p. 63(

† *Ibid.* p. 614.

has the same association coefficient as

$$\begin{array}{c|c} a & b \\ \hline c & d \end{array}:$$

and that accordingly if the coefficient of association were a valid measure of relationship every fourfold could be expressed in a form which led to a Gaussian

$$r = \sin\frac{\pi}{2}\left(\frac{\sqrt{ad}-\sqrt{bc}}{\sqrt{ad}+\sqrt{bc}}\right) = Q_3.$$

Thus Pearson's Q_3 was a direct result of writing the fourfold in the "equalised" form

$$\begin{array}{c|c} \sqrt{ad} & \sqrt{bc} \\ \hline \sqrt{bc} & \sqrt{ad} \end{array}$$

which Mr Yule now proposes as a primary virtue of his method as giving all classes their "natural" or equal percentages. The new "coefficient of colligation" is thus really an old friend, which under the form Q_3 did not possess "the fundamentally different properties*" with which Mr Yule credits it, being the direct and we venture to think legitimate offspring of the equalised frequency table as figured above.

The next historical point where Mr Yule seems to be at fault—at any rate in his criticisms of one of the present writers—is in his interpretation of the word *correlation.* He narrows it down to the significance of correlation coefficient found by the product-moment formula, and so obsessed is he by this idea that he applies it to a correlation of gross ranks, which is not a correlation of variates at all. The word correlation in the statistical as distinguished from the biological sense, we believe, was first used by Galton in his memoir of 1889 entitled: "Co-relations and their Measurements, chiefly from Anthropometric Data†," and he gave a definition of it which does not involve the conception of the product moment or the linearity of regression at all. That definition was extended by one of the present writers in a memoir of 1895‡—and it runs: "Two organs in the same individual, or in a connected pair of individuals, are said to be correlated, when a series of the first organ of a definite size being selected, the mean of the sizes of the corresponding second organs is found to be a function of the size of the selected first organ. If the mean is independent of this size, the organs are said to be non-correlated. Correlation is defined mathematically by any constant, or series of constants, which determine the above function." It will be seen that this definition of correlation has nothing whatever in it that limits the use of the word 'correlation' to the coefficient of correlation as found by the product-moment method. Galton himself never used the product-moment method to find his "index of correlation." He had generally in view§ the position that the average value of one organ or

* *Journn. of R. S. S.* Vol. LXXV. p. 592, footnote.

† *Royal Soc. Proc.* Vol. XLV. p. 135.

‡ [illegible] A, p. 257.

§ [illegible] id to be co-related when the variation of the one is accompanied on [illegible] tion of the other...." *R. S. Proc.* Vol. XLV. p. 135.

attribute associated with a given value of a second changes continuously as we change that given value. It was only his wide experience of anthropometric data which led him to believe that in most cases the function that the mean of one organ is for a given value of the second may be adequately represented by a straight line. There was nothing either in his own treatment or in the work of his followers of the Biometric School, which pinned them down to the use of the word correlation for a particular constant found by the product-moment method. The "correlation ratio" has in the work of that school just as much significance as the "correlation coefficient," and it is only Mr Yule who proposes to confine the use of the word correlation to the narrow sense of straight line regression determined by product-moment methods. To the biometrician correlation when it ceases to be linear is not determined by the product-moment value of r at all, and the grade of correlation may be far higher than the value as determined by the coefficient of correlation. In the same way other constants may be found defining the relationship of two continuous characters, or measuring their degree of dependence. These are equally measures of correlation in our sense of the word. Mr Yule's coefficient of association is not in our sense of the word a measure of correlation at all, it shows in no manner how the mean of one attribute for a given value of a second attribute varies as we modify this value. It is, as we shall show below, impossible to give it in the case of continuous variates any rational significance whatever. Where there is no true correlation at all, the size of Mr Yule's Q may be produced solely by a lack of homoscedasticity—of equal variation—in the arrays of one variable associated with constant values of the second, but in what manner it measures this heteroscedastic property is quite beyond interpretation. Mr Yule claims that the nature of the frequency is of no consequence, he states that the coefficient of association may be applied without any general theory of frequency. For us this is not a correct attitude; we admit wide deviations from Gaussian distributions, but such cases are not the rule. Mr Yule can pick out special instances which are far from Gaussian, such as age distributions, barometric heights, or heterogeneous mixtures of *growing* organs like ivy leaf lengths. Even for such cases he has not examined in any adequate manner how far methods based on Gaussian or other allied considerations do give practical results, nor how far even the Gaussian fourfold r—tetrachoric r or r_t, we will call it for the purposes of this paper—is a more stable and reliable coefficient than those suggested by himself.

He has indeed criticised the application of a tetrachoric r_t to eye-colour data—his discussion of the subject will be considered in a separate section of our reply—but he does not inform his readers that the one of the present writers concerned in the eye-colour treatment fully admitted in June 1907* "the unsatisfactory approach to the Gaussian distribution found in pigmentation tables" and stated that his very knowledge of this point had led him to develop the method of contingency for such problems. Yet Mr Yule raises *five* years later the applicability of the

* *Biometrika*, Vol. v. p. 472.

normal coefficient, the tetrachoric r_t, to pigmentation tables without a line of reference to a paper, which must have been perfectly well known to him, for it is entitled "A Reply to Certain Criticisms of Mr G. U. Yule." This may be strategic, but at the same time illustrates the peculiar character of Mr Yule's controversial methods.

(3) *On the Boas-Yulean "Theoretical Value" of the Correlation (Pearson's ϕ).*

We now come to a very nice point indeed, namely Mr Yule's "theoretical value of the correlation coefficient," the method introduced into his Textbook of Statistics (p. 212) without a word of warning as to when it should be applied. Mr Yule now states that: "It would have been thought that anyone reasonably acquainted with the theoretical work of the last decade, and especially Professor Pearson and his collaborators, would have found no difficulty in the passage in question*." Now what exactly have Pearson and his collaborators done? They have applied the product-moment correlation to the presence of 0, 1, or 2 protogenic *units* in *theoretical* Mendelian investigations. They have assumed that when a character goes by *units*, you may apply the usual product-moment methods. But they have objected *in toto* to the application of such a method to material where there was reasonable evidence of continuous variation. Does Mr Yule look upon 'death' as the addition of one unit to 'recovery'? Does Mr Yule look upon vaccination as the addition of one unit to 'absence of vaccination'? Does Mr Yule look upon 'mental defect' as the addition of one unit to normal mentality? In the three Mendelian types (RR), (DR), and (DD) there is a progression of one unit at each stage in the number of D's, but what are the units in the cases we have mentioned? Or, does Mr Yule suggest that his "theoretical value of the correlation" is to be confined to those actual true unit additions to which Pearson and his collaborators have always confined them? There is not a hint of this in his Textbook nor in his present paper. He has indeed carefully refrained hitherto from saying what are the characters of the attributes to which it is to be applied. He has suggested that "the ordinary theory of correlation, once that theory had been freed from any necessary relation to the theory of normal correlation, was applicable in its entirety to the 2×2-fold table†." Mr Yule says that this should be "a very obvious matter." Indeed!—Then apparently 'vaccination' and 'mental defect' are a quantitative unit more than non-vaccination and normal mentality!—But how does this fit with Mr Yule's other assertion that these are discrete attributes and suitable for the application of his coefficient of association? Let us see how Dr Boas investigates Pearson's $r_{hk} = \phi$‡.

"Correlations of phenomena that cannot be measured but only counted may be treated in the following manner: If two events that have the probabilities p_1 and p_2 are correlated, we may say that those cases in which the event 1 occurs

* *Journal of R. S. S.* Vol. XLV. p. 609.

† *Journal of R. S. S.* Vol. XLV. p. 606.

‡ *Science*, May 1, 1909, p. 824.

have the probability 1, or a deviation from the normal probability $1-p$. Those cases in which the event 1 does not occur have the probability 0, or the deviation from the average probability of $-p_1$."

Pearson at the time read this many times through; both of the present writers have read and re-read it since, and they fail utterly to grasp how an event can have at the same time a probability of 1 and of 0 and a normal probability of $1-p$! Mr Yule says that Professor Pearson failed "to understand what Dr Boas was doing*." We still fail completely to understand what Dr Boas means, or how Mr Yule justifies the assertion that Dr Boas was demonstrating a formula which applied to two values of a character differing by a unit. What did come out of Dr Boas' investigation of 1909 when it was translated into the fourfold table terminology

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	N

was
$$r_{hk}=\frac{ad-bc}{\sqrt{(b+d)(a+c)(c+d)(a+b)}}=\phi,$$

a value already known (i) as the correlation r_{hk} between the means, *each measured in terms of their standard deviations*, of two variates of a fourfold Gaussian table† or (ii) as the square root of the mean square contingency of a fourfold table without any assumption of a Gaussian distribution‡.

Now it is known that the correlation of errors in two means, $r_{\bar{x}\bar{y}}$, is equal to the correlation of deviations, r_{xy}, in the two variates of which $\bar{x}$ and $\bar{y}$ are the means. But r_{hk} is *not* $r_{\bar{x}\bar{y}}$, for $h=\bar{x}/\sigma_x$ and $k=\bar{y}/\sigma_y$ and σ_x and σ_y are correlated as well as $\bar{x}$ and $\bar{y}$. It is therefore clear that if x and y are continuous variables of any kind, r_{hk} is not a "theoretical value of the correlation" of variates. It is a correlation of ranks where the ranking consists of only first and second, and is wholly uncorrected for class-index. It becomes a true value of the correlation when the two classes differ by a unit quantity as in the units of *theoretical* Mendelism. There is not a word of this in Dr Boas' paper; he speaks of his formula as applicable when things can be counted but not measured. Mr Yule speaks of it as applicable in its entirety to the 2×2-fold table!—When pressed he says it would have been thought that no one acquainted with the work of the Biometric School—on Mendelism—could fail to understand what his passage signified. We wholly fail to understand it now. Is r_{hk}, i.e. ϕ, applicable to every fourfold division? If so, why does not Mr Yule use it and drop his coefficient of association?—In truth he

* *Journal of R. S. S.* Vol. XLV. p. 608.

† *Phil. Trans.* Vol. 195 A (1900), p. 12.

‡ *Drapers' Company Research Memoirs*, "On the Theory of Contingency," 1904, p. 21.

cannot do it without confessing himself hopelessly in error!*—If ϕ or r_{hk} is the right thing to use for a fourfold table as Mr Yule now suggests, then his ω and his Q are hopelessly wrong, for all the selections which do not alter his Q—and the fact that it is not altered by selection is according to him one of its great merits—do alter, and this to any extent we please, his "theoretical value of the correlation." This very point was emphasised by one of the present writers in a paper recently published in *Biometrika*† but Mr Yule appears wholly to have missed the essential features of that criticism. The points were that (i) there was a wide range of values of Q—Mr Yule's coefficient of association—for a surface of the same correlation; if it is impossible to compare the values of Q for the same surface divided at different places with any intelligible result, what possible comparison can be made of Q from one system to a second?, and (ii) the values of Q and r_{hk} or ϕ are wholly different and tend in opposite directions as we change our divisions. Under "wholly different" we include liability to be "wholly differently interpreted." Both Q and ϕ range numerically between 0 and 1. Therefore in estimating the meaning of Q and ϕ or r_{hk} we have to consider where they stand on this range. Mr Yule examining the associations between developmental defects, nerve signs, low nutrition and mental dulness finds values ranging from ·75 to ·95 for his coefficient Q. He comments: "The associations are, however, all high (very high compared with most coefficients of organic correlation with which one has to deal‡), ranging from ·784 (? ·750) to ·952§." Elsewhere in the same paper Mr Yule speaks of ·174 as a "very small association," and a ·8 to ·9 association "as very high indeed ||." We know accordingly what Mr Yule understands by high and low association. Indeed if a scale of values is to lie between 0 and 1, those approaching 0 must be very low and those approaching 1 must be very high. Now Heron applied Mr Yule's or Dr Boas' "theoretical coefficient" to precisely the same data as those for which Mr Yule had calculated his association Q, and found that it was very high. Heron found that for $Q=·921$ and ·753, the "theoretical value of r" $=·011$ and ·006 respectively. If both these ways of investigating relationship are valid, then ·011 and ·006 must on a correlation scale represent a "very high degree of association." It would be interesting to know how Mr Yule would describe $\phi=·95$ or what represents a *low* association, if ·01 corresponds to a high degree of association!—But any one who is familiar with coefficients of correlation—and ϕ or r_{hk} is a real coefficient of correlation,—knows that values of ·01 and under are extremely low values and, whatever their probable errors may be, are of no significance for purposes of prediction. All Mr Yule can say in reply to Heron's statement that one of Mr Yule's methods gives very high relationship and the

* See also pp. 172–4 below, where this point is touched on again.

† "The Danger of Certain Formulae suggested as substitutes for the Correlation Coefficient," by David Heron, Vol. VIII. p. 109.

‡ Here and at other points of his earlier papers, Mr Yule apparently considers that Q is really comparable with the true correlation.

§ *Phil. Trans.* Vol. 194 A, p. 300. See also "high degree of association," *Theory of Statistics*, p. 34.

|| *Ibid.* pp. 289 and 296.

other very low relationship between the attributes under investigation, is that "these small figures represent perfectly appreciable intensities of the product sum correlation r." But this is an entirely different matter*, the intensity of correlation on the scale 0 to 1 is quite distinct from its degree of reliability. Mr Yule would have termed ·95 a very high association had it been deduced from 1000 cases, and we term ·01 a very low correlation even when deduced from a census population. Here are the results for which we have calculated the probable errors†.

* Mr Yule himself says (*loc. cit.* p. 651) that the statistician must always keep apart the magnitude of the association and the reliability of this magnitude. We doubt the truth of this statement, but it cuts against his own argument at this point!

† The probable error of a coefficient of correlation for *any* frequency distribution is given by the formula:

$$\text{P.E. of } r = \cdot 67449\sigma_r = \frac{\cdot 67449 r}{\sqrt{n}}\left\{\frac{p_{22}}{p_{11}^2}+\frac{p_{22}}{2p_{20}p_{02}}+\frac{p_{40}}{4p_{20}^2}+\frac{p_{04}}{4p_{02}^2}-\frac{p_{31}}{p_{11}p_{20}}-\frac{p_{13}}{p_{11}p_{02}}\right\}^{\frac{1}{2}}$$

(Pearson, *Drapers' Research Memoirs*, "General Theory of Skew Correlation," p. 20).

The whole work therefore of finding σ_r turns on evaluating p_{11}, p_{22}, p_{31}, p_{13}, p_{20} and p_{02} for a fourfold table:

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	n

when we suppose *concentration at single points* of the frequencies.

We write $$ad-bc=\epsilon, \quad (ad-bc)/n^2=\epsilon',$$
$$(a+b)(c+d)(a+c)(b+d)=q, \quad \text{and} \quad q/n^4=q'.$$

Then we have $$r=\epsilon'/\sqrt{q'},$$

$$p_{11}=\epsilon', \qquad p_{22}=q'+\epsilon'\left(\frac{c+d}{n}-\frac{a+b}{n}\right)\left(\frac{b+d}{n}-\frac{a+c}{n}\right),$$

$$p_{20}=\frac{(b+d)(a+c)}{n^2}, \qquad p_{02}=\frac{(a+b)(c+d)}{n^2},$$

$$p_{31}=\epsilon'\frac{(b+d)^3+(a+c)^3}{n^3}, \qquad p_{13}=\epsilon'\frac{(c+d)^3+(a+b)^3}{n^3}.$$

Write for brevity $$m_1=\frac{c+d}{n}, \qquad m_1'=\frac{a+b}{n},$$
$$m_2=\frac{b+d}{n}, \qquad m_2'=\frac{a+c}{n},$$

and $$\phi'=\frac{1}{m_1}+\frac{1}{m_2}+\frac{1}{m_1'}+\frac{1}{m_2'}.$$

Then $$\psi=\frac{(b+d)^2}{a+c}+\frac{(a+c)^2}{b+d}+\frac{(c+d)^2}{a+b}+\frac{(a+b)^2}{c+d}=n(\phi'-6).$$

And $$\frac{p_{22}}{p_{11}^2}=q'/\epsilon'^2+\frac{(m_1-m_1')(m_2-m_2')}{\epsilon'},$$

$$\frac{p_{22}}{2p_{20}p_{02}}=\frac{1}{2}+\frac{1}{2}\frac{\epsilon'}{q'}(m_1-m_1')(m_2-m_2'),$$

$$\frac{p_{40}}{4p_{20}^2}+\frac{p_{04}}{4p_{02}^2}=\frac{1}{4n}\psi=\frac{1}{4}(\phi'-6),$$

$$\frac{p_{31}}{p_{11}p_{20}}+\frac{p_{13}}{p_{11}p_{02}}=\frac{\psi}{n}=\phi'-6.$$

Relationship between Blindness and Mental Defect for different Age Groups.

Age Group	"Theoretical value of r" or ϕ
5—	$\cdot 0113 \pm \cdot 0030$
10—	$\cdot 0100 \pm \cdot 0022$
15—	$\cdot 0065 \pm \cdot 0017$
20—	$\cdot 0061 \pm \cdot 0015$
25—	$\cdot 0046 \pm \cdot 0009$
35—	$\cdot 0060 \pm \cdot 0010$
45—	$\cdot 0053 \pm \cdot 0008$
55—	$\cdot 0059 \pm \cdot 0012$
65—	$\cdot 0028 \pm \cdot 0012$
75—	$- \cdot 0031 \pm \cdot 0014$
85—	$\cdot 0058 \pm \cdot 0065$
All ages *5—85*	$\cdot 0066 \pm \cdot 0002$

Is it conceivable that, if Mr Yule had approached the problem of the relationship of blindness and mental defect from the standpoint of "the theoretical value" of r and found that the maximum value of the coefficients obtained was ·0113, there would have been any talk of the high association of the two attributes? If this on the Boas-Yulean scale means high association, what language would Mr Yule find to describe a Boas-Yulean coefficient of ·96? Instead of replying to this criticism of Heron, Mr Yule states that this series of values confirms his view that the association decreases with age! If the reader looks at the diagram below, he will see the mean value of the Boas-Yulean coefficient with twice the probable error of each age sample set off either side of it. He will note (i) that there is only one age 75—80 where the deviation from the average value becomes significant; (ii) that from age 15 to age 60, the polygon is practically horizontal and agrees with the mean; (iii) the "high" values (·01 order!) occur in childhood, especially ages 5—10, where diagnosis of mental defect is doubtful, and the low values at 65 and onwards (where they are even contradicted by age 85 onwards!), just the ages when senile decay and old age cataract may lead the recorder of a census return to almost any statement as to mental derangement and blindness. We feel fairly confident that the unbiased statistician with this result before him could only con-

Hence $$\sigma_r^2 = \frac{1}{n}\left\{1 + \frac{\epsilon'}{q'}(m_1 - m_1')(m_2 - m_2') - (\tfrac{3}{4}\phi' - 5)\frac{\epsilon'^2}{q'} + \tfrac{1}{2}\frac{\epsilon'^3}{q'^2}(m_1 - m_1')(m_2 - m_2')\right\}$$

$$= \frac{1}{n}\{1 - r^2 + (r + \tfrac{1}{2}r^3)\lambda\mu - \tfrac{3}{4}r^2(\lambda^2 + \mu^2)\},$$

or $$\text{P.E. of } r = \frac{\cdot 67449}{\sqrt{n}}\{1 - r^2 + (r + \tfrac{1}{2}r^3)\lambda\mu - \tfrac{3}{4}r^2(\lambda^2 + \mu^2)\}^{\frac{1}{2}},$$

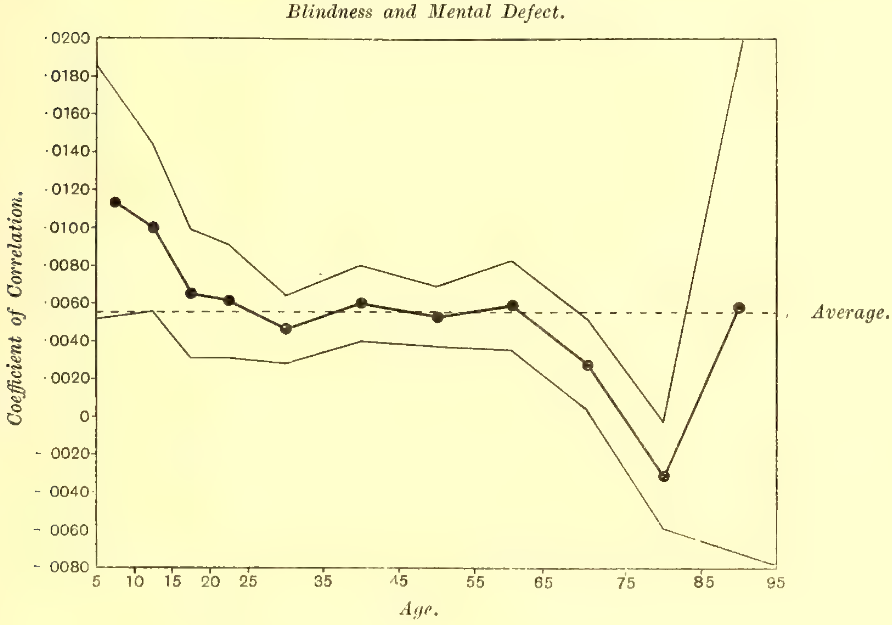
where $$\lambda = \sqrt{\frac{m_1}{m_1'}} - \sqrt{\frac{m_1'}{m_1}} \text{ and } \mu = \sqrt{\frac{m_2}{m_2'}} - \sqrt{\frac{m_2'}{m_2}}.$$

This form agrees with that obtained by Yule and verified by Greenwood, but our deduction of it appears to be the natural method and shows its relation to the general formula for the probable error of r.

clude that the Boas-Yulean method indicated no practically important relationship between mental derangement and blindness and that there was no *trustworthy* sign of any modification of this relationship with age. The point is not whether such a relationship really exists or not, but that one method advocated by Mr Yule shows definite results (high association), although a second, equally strongly advocated, fails to give any association of practical value and is thus in direct opposition to the first*. Mr Yule has omitted to indicate which method is the proper one to use in such cases. Are both right or both wrong? One method Mr Yule states may be applied to all 2 × 2-fold tables, the other should be used for "discrete" quantities. When is a quantity "discrete"? Mr Yule confuses "discreteness" in the class-index—a mere verbalism—with discreteness in the attribute classified under it, and this reduces his investigations from the plane of practical statistics to the field where we originally placed them, that of theoretical logic.

DIAGRAM I. Diagram showing absence of any relationship of practical value between Blindness and Mental Defect.

Blindness and Mental Defect.



(4) *On Association and the Boas-Yulean Coefficient.*

Again we reach an interesting point which Mr Yule has failed to elucidate. It is best illustrated by an example. Take the following table for lengths of ivy leaves on the same spray—one of the examples selected by Mr Yule:

* See also p. 204 below.

A. Length of First Leaf.

Length of Second Leaf.		Under *6·95* mm.	Over *6·95* mm.	Totals
	Under *14·95* mm. ...	6,943	46,673	53,616
	Over *14·95* mm. ...	41	6,343	6,384
	Totals	6,984	53,016	60,000

The true correlation of the material is ·5672 and the tetrachoric $r = \cdot 5572 \pm \cdot 0058$. It may be presented in any one of the three forms (i) as A above, (ii) in one of Mr Yule's "natural forms *," e.g.

B.

6,636	1,383	8,019
1,383	6,636	8,019
8,019	8,019	16,038

Or, again, as

C.

6,943	4·6673	6,947·6673
410,000	6,343	416,343
416,943	6,347·6673	423,290·6673

Now Q is the same = ·9167, "very high association" for all these forms, and *a posteriori* given Q we should not know whether it had arisen from A, B or C. The values of the Boas-Yulean coefficient for these tables are:

A. $\phi = \cdot 1183$,
B. $\phi = \cdot 6551$,
C. $\phi = \cdot 0152$.

Which of these values are we to adopt? Are we to use the table as it originally is given? Or the "natural" form B? Or the selected form C? How can two coefficients ever lead to the same results, when an adjusting process which does not modify one, a property claimed as one of its advantages, changes the other to any value we please from zero up to Mr Yule's new coefficient of colligation†?

* See *Journal R. Stat. Soc.* Vol. LXXV. p. 590.

† Taking the usual form of the fourfold:

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	n

Let us examine the effect of "adjusting" the tables in Mr Yule's cases of vaccination and death at Sheffield, Leicester and Homerton-Fulham. He asserts that it is "natural" to take 50 °/₀ of vaccinated; we fail to understand why 50 °/₀ is more or less "natural" than 70 °/₀ or 95 °/₀ or than the percentage which actually occurs in the smallpox cases of those towns. Even if he takes 50 °/₀ of vaccinated, why it should be "natural" to take 50 °/₀ of deaths also is to us equally mysterious, and we believe must be to that juryman "any man of ordinary intelligence" to whom Mr Yule appeals. The following table gives the values that would arise from different methods of "adjusting" the tables:

Smallpox—Vaccination and Death.

	Yule's Association Q	Boas-Yulean ϕ				r_t	C_2	P
	(a)	(b)	(c)	(d)	(e)			
Percentage of Deaths ...	50 °/₀	50 °/₀	Actual	Actual	Actual	Actual	Actual	Actual
Percentage of Vaccinations	50 °/₀	50 °/₀	70 °/₀	Actual	95 °/₀			
Sheffield	·902	·630	·531	·479	·383	·769	·432	$5{\cdot}15/10^{234}$
Leicester	·862	·572	·249	·233	·190	·611	·187	$2{\cdot}46/10^{4}$
Homerton-Fulham ...	·804	·504	·423	·409	·084	·662	·379	$1{\cdot}22/10^{414}$

It is clear that the Boas-Yulean method will give any results whatever between zero and those in the (b) column, according to what percentage we choose to take in the adjusted table of deaths and vaccinations. We can also change the perfectly arbitrary order that Mr Yule has given for the three towns. It appears to us that his statement that "it should have been an obvious matter that the ordinary theory of correlation, once that theory had been freed from any necessary relation to the theory of normal correlation, was applicable in its

modify it by a Yulean selection using the factors l and m into

la	b	$la+b$
lmc	md	$m(lc+d)$
$l(a+mc)$	$b+md$	—

Q remains unchanged. But ϕ now takes the value

$$\frac{lm(ad-bc)}{\sqrt{lm(a+mc)(b+md)(la+b)(lc+d)}}=(ad-bc)\Big/\sqrt{\left(ad+bc+\frac{ab}{m}+mcd\right)\left(ad+bc+\frac{bd}{l}+lac\right)}.$$

This is a minimum when l and m are indefinitely great; it is a maximum when

$$m=\sqrt{ab}/\sqrt{cd}, \qquad l=\sqrt{bd}/\sqrt{ac},$$

or

$$\phi=\frac{\sqrt{ad}-\sqrt{bc}}{\sqrt{ad}+\sqrt{bc}}=\omega, \text{ the coefficient of colligation.}$$

entirety to the 2×2-fold table" is wholly inconsistent with any validity in the coefficient of association.

It is not only that we can give a vast range of values to ϕ for a constant Q, but equally we can give Q a whole range of values, starting at the symmetrical table value and proceeding up to unity for a constant ϕ. Examine for example the following series of fourfold tables. The first is the table as it actually occurred with $Q = \cdot 5078$, the second is Mr Yule's "equivalent" symmetrical table with of course the same Q.

(1)

29600	17300	46900
37200	66600	103800
66800	83900	150700

$Q = \cdot 5078, \quad \phi = \cdot 2542.$

(2)

47952	27398	75350
27398	47952	75350
75350	75350	150700

$Q = \cdot 5078, \quad \phi = \cdot 2728.$

We now proceed to adjust the latter table so that ϕ remains stationary and Q rises.

(3)

45452	47103	92555
12693	45452	58145
58145	92555	150700

$Q = \cdot 5511, \quad \phi = \cdot 2728.$

(4)

43052	56385	99437
8211	43052	51263
51263	99437	150700

$Q = \cdot 6003, \quad \phi = \cdot 2728.$

(5)

39852	66467	106319
4529	39852	44381
44381	106319	150700

$Q = \cdot 6813, \quad \phi = \cdot 2728.$

(6)

37952	71808	109760
2988	37952	40940
40940	109760	150700

$Q = \cdot 7407, \quad \phi = \cdot 2728.$

(7)

35852	77349	113201
1647	35852	37499
37499	113201	150700

$Q = \cdot 8196, \quad \phi = \cdot 2728.$

(8)

33552	83090	116642
506	33552	34058
34058	116642	150700

$Q = \cdot 9288, \quad \phi = \cdot 2728.$

(9)

32952	84541	117493
255	32952	33207
33207	117493	150700

$Q = \cdot 9611, \quad \phi = \cdot 2728.$

(10)

32552	85499	118051
97	32552	32649
32649	118051	150700

$Q = \cdot 9845, \quad \phi = \cdot 2728.$

(11)

32352	85974	118326
22	32352	32374
32374	118326	150700

$Q = \cdot 9964, \quad \phi = \cdot 2728.$

(12)

32327	86035	118362
11	32327	32338
32338	118362	150700

$Q = \cdot 9982, \quad \phi = \cdot 2728.$

(13)

32302	86095	118397
1	32302	32303
32303	118397	150700

$Q = \cdot 9998, \quad \phi = \cdot 2728.$

(14)

32298	86104	118402
0	32298	32298
32298	118402	150700

$Q = 1 \cdot 0000, \quad \phi = \cdot 2728.$

The two series of results just considered show us that with ϕ constant we can make Q rise to unity and with Q constant we can make ϕ fall to zero. It is therefore always possible to determine or to find in actual practice series of tables in which an ascending order of Q is accompanied by a descending order of ϕ. That is to say the two coefficients will flatly contradict each other. There is no basis whatever for Mr Yule's assertion that in Sheffield the correlation of vaccination and recovery is highest, while in Homerton and Fulham it is lowest. With the same three values of Q that he gives, we can make any order of the ϕ's we please. In the actual tables the relative order is not that given by Mr Yule; the order of tetrachoric r agrees with that of ϕ; so does the coefficient of mean square contingency and further the probability of independence, i.e. the probability that death is independent of vaccination. These are given as r_t, C_2 and P in the last three columns of the table on p. 173, and they entirely reverse Mr Yule's judgment. P is of course in a different class to any other of the coefficients, but we return to this point later.

Mr Yule states that any table may be reduced without modification of the association to its equalised form; for example the tables

998,667	666	999,333
666	1	667
999,333	667	1,000,000

and

300,000	200,000	500,000
200,000	300,000	500,000
500,000	500,000	1,000,000

are "equivalent," both have $Q = \cdot 385$, but the inferences to be drawn from the two tables had they *originated independently* are quite different. In the first case the probable error of $Q = \cdot 288$, and the result is not definitely significant. In the second case the probable error is $\cdot 001$ and there is no doubt of the significance. If Q_O be the Q of an original table and Q_E of the equalised table, then we have Mr Yule's vaccination data:

	P. E. of Q_O	P. E. of Q_E
Sheffield	·007	·003
Leicester	·065	·022
Homerton and Fulham	·007	·005

Mr Yule has not entered into this question of the probable errors of the series of his modified tables. The statistician, however, whose long experience enables him closely to associate given types of tables with given degrees of reliability, is largely deceived when an "equivalent table" is presented to him of which the probable error as it stands may be $\frac{1}{2}$ to $\frac{1}{100}$ or less of the true probable error of the coefficient given.

We have shown in this section of our discussion that the two coefficients Q and ϕ cannot both be valid. Mr Yule nowhere adequately assigns the type of cases to which the one or the other should be applied. He tells us that he hopes the

"comedy of errors" has now ended, that comedy consisting in our overlooking the "fact" that it is "a very obvious matter" that ϕ is applicable in its entirety to the 2×2-fold table. That applicability is the very point to which Pearson took exception in Dr Boas' use of ϕ. Were it a fact, Mr Yule might throw both his Q and his ω overboard, for what more is required than a method applicable in its entirety to a 2×2-fold table? And the ϕ method at all points contradicts the Yulean results. But it is not true, for at best it would be a method of ranks, and the correlation of ranks is never a correlation of variates unless the ranked quantities proceed by absolute units of variation—as for example in the *theoretical* Mendelian case to which Pearson perfectly legitimately applied it—or in counting the teeth on the carapace of a prawn or the veins on a leaf.

Mr Yule seems to consider that the inoculation with an antitoxin is equivalent to the addition of a unit of something to the individual; we consider that this is wholly erroneous. To begin with, the dosage is not uniform, its repetition does not always occur at the same interval and the number of doses is not always the same; further the interval between the onset of the disease and first inoculation is by no means the same; lastly, apart from the resistance of the individual patients to the disease, the curative effect of the treatment depends on the relation of the antitoxin administered to the physiological individuality of the patient. It is idle therefore to consider this varying complex as a quantity undifferentiated from individual to individual. The group treated with antitoxin is not made up of identical individuals but of a number of persons with increased power of resistance to the disease, which may vary from the case of a person who has gained nothing by it to that of a person who has immensely increased his power of recovery. In precisely the same way those who have not been treated can by no means be grouped into a single quantitative class; it may be doubted, indeed, whether recuperative power when disease is incurred is really divided sharply by a line like treatment or non-treatment with antitoxin. It may be only the sharpest division we can take under the circumstances, and in our ignorance of the nature of the distribution a tetrachoric r_t may be as effective a measure of association as Mr Yule's Q. At any rate no man of "ordinary intelligence" would believe that perfect association existed between treatment and recovery because out of 23 persons treated none died, while out of 977 not treated six died, yet this would be the result provided by Mr Yule's coefficient of association! Clearly if only $0{\cdot}6\,^\circ/_\circ$ died without treatment, we should not expect *any* to die in a sample of 23 whether treated or not treated. The vanishing of association for a zero quadrant is a patent fallacy*.

If the problem were presented to us as Mr Yule states it, i.e. the evaluation of a new treatment, we should certainly not use his Q for solving the problem. We should probably to-day not use a tetrachoric r_t, except as a control. We should most likely use ϕ, although certainly not in the sense of a correlation of four points. We should question how far death and recovery are independent of treatment or non-treatment; that is to say, we should ask what is the probability that recovery

* This topic is of such importance that we have discussed it more at length in Appendix I.

was independent of treatment. If n be the total number of observations, then, if $n\phi^2 = \chi^2$ be calculated, the chance of independence is at once given by finding the P corresponding to χ^2 from the tables for "goodness of fit" in the case of $n = 4$. The column of P in the table of p. 173 shows us that the Homerton-Fulham data stand at the head of the series in this respect. This is of course in great part due to the larger numbers dealt with, but obviously in such a question as the effectiveness of treatment weight must be given to numbers.

If we turn to the problem of r_t and C_2 considered merely as coefficients of association, we must examine what Mr Yule has laid down as the desiderata of such a coefficient. The fourfold table being $\frac{a \mid b}{c \mid d}$, he has *assumed* that the coefficient must range from -1 to $+1$, and that if any *one* of the cells be empty the coefficient must be $+1$ or -1. He then guesses a formula $\frac{ad - bc}{ad + bc}$ or another $\frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}$ out of the many thousands which can be invented*.

But is either of the assumptions above really necessary? Why should the association be perfect when $b = 0$? Why should it be perfect even if both b and c are zero? Let us toss a shilling and a penny and record heads or tails of both.

Shilling.

Penny.		Head	Tail	Totals
	Head ...	1	0	1
	Tail ...	2	1	3
	Totals ...	3	1	4

We do it four times and the result is as above, on the whole not a very improbable result. But according to Mr Yule there is absolute association, and since the probable error according to him is zero, the result is *absolutely reliable*. Clearly

* See Pearson, *Phil. Trans.* Vol. 195 A, p. 15. If $\phi(z)$ be a function which vanishes with z, then any form

$$\{\phi(ad) - \phi(bc)\}/\{\phi(ad) + \phi(bc)\}$$

satisfies Yule's requirements. Or, we can take a form

$$Q_\kappa = \frac{1-\kappa}{1+\kappa} \frac{\phi\left(\kappa + \frac{1}{\kappa}\right)}{\phi(\infty)}$$

if $\phi(\kappa)$ be finite for $\kappa = \infty$, where $\kappa = (bc)/(ad)$. Clearly for the range 0 to ∞ of values of κ, Q_κ ranges from $+1$ to -1 and satisfies Mr Yule's conditions if $\phi(\infty)$ is $> \phi(\kappa)$, but by an arbitrary choice of ϕ we can get any form of Q-curve we please. No curve of real significance can be obtained, i.e. no reasonable value of an association coefficient by the simple condition of fixing three of its values without other hypothesis!

ϕ would be better than this, for it is equal to ·33 with a probable error of ·15 which makes it for practical purposes unreliable. Or again, we repeat our experiment and find

Shilling.

		Head	Tail	Totals
Penny.	Head ...	2	0	2
	Tail ...	0	2	2
	Totals ...	2	2	4

Here Q is again unity or there is absolute association, and the probable error is zero or this association according to Mr Yule is also absolutely reliable. Further $\phi = 1$, and its apparent probable error is zero, but this is only apparent because the calculation of the probable error (as indeed of that of Mr Yule's Q, although he has not noticed it) is incorrect for such a case.

These simple illustrations seem to indicate that there is nothing in the nature of things which necessarily demands that the association shall be unity when either b or c or both are zero. On the other hand the probability P (as derived from ϕ, the mean square contingency) that the heads and tails of shilling and penny are independent is more than ·90, and in the second case more than ·25. It seems to us therefore that this mean square contingency method which gives reasonably satisfactory results, where the Boas-Yulean goes hopelessly astray, is far more likely to be preferable in the case of medical treatment to which Mr Yule proposes to apply his coefficient.

There is another point also to be considered. Why should the range of a good coefficient of association lie for any given number of cells between $+1$ and -1? Let us examine the following table:

(A)

	a	b	c	d	e	f	g	h	i	j	Totals
α	40	0	0	0	0	0	0	0	0	0	40
β	0	0	200	0	0	10	110	10	30	0	360
γ	0	0	0	0	0	10	440	0	120	0	570
δ	0	0	0	0	0	10	20	0	0	0	30
ϵ	0	0	0	20	10	0	0	0	0	0	30
ζ	0	120	0	440	10	0	0	0	0	0	570
η	0	30	10	110	10	0	0	200	0	0	360
θ	0	0	0	0	0	0	0	0	0	40	40
Totals	40	150	210	570	30	30	570	210	150	40	2000

If we treat each sub-range here as unity we find the correlation *negative* and equal to $-$ ·1120. If therefore we assume this to be the ultimate distribution, this is the correlation coefficient.

Now combine b and c, h and i. Proceeding in the same way to find the Yulean pseudo-ranks r, we have now for the following table:

(B)

	a	$b+c$	d	e	f	g	$h+i$	j	Totals
α	40	0	0	0	0	0	0	0	40
β	0	200	0	0	10	110	40	0	360
γ	0	0	0	0	10	440	120	0	570
δ	0	0	0	0	10	20	0	0	30
ϵ	0	0	20	10	0	0	0	0	30
ζ	0	120	440	10	0	0	0	0	570
η	0	40	110	10	0	0	200	0	360
θ	0	0	0	0	0	0	0	40	40
Totals	40	360	570	30	30	570	360	40	2000

the value is positive and equal to + ·0050.

Now club d and e, f and g, γ and δ, ϵ and ζ together, and we have

(C)

	a	$b+c$	$d+e$	$f+g$	$h+i$	j	Totals
α	40	0	0	0	0	0	40
β	0	200	0	120	40	0	360
$\gamma+\delta$	0	0	0	480	120	0	600
$\epsilon+\zeta$	0	120	480	0	0	0	600
η	0	40	120	0	200	0	360
θ	0	0	0	0	0	40	40
Totals	40	360	600	600	360	40	2000

The Yulean pseudo-ranks r is now + ·2562.

Combine a and $b+c$, $h+i$ and j, α and β, η and θ, and we find

(D)

	$a+b+c$	$d+e$	$f+g$	$h+i+j$	Totals
$\alpha+\beta$	240	0	120	40	400
$\gamma+\delta$	0	0	480	120	600
$\epsilon+\zeta$	120	480	0	0	600
$\eta+\theta$	40	120	0	240	400
Totals	400	600	600	400	2000

The Yulean now drops down to + ·1429!

Combine $d+e$ and $f+g$, $\gamma+\delta$ and $\epsilon+\zeta$ to give

(E)

	$a+b+c$	$d+e+f+g$	$h+i+j$	Totals
$\alpha+\beta$	240	120	40	400
$\gamma+\delta+\epsilon+\zeta$	120	960	120	1200
$\eta+\theta$	40	120	240	400
Totals	400	1200	400	2000

The Yulean is now + ·5000.

Combine $b+c$, $d+e$, $f+g$, $h+i$ together; β, $\gamma+\delta$, $\epsilon+\zeta$, η, and we deduce

(F)

	a	$b+c+d+e+f+g+h+i$	j	Totals
a	40	0	0	40
$\beta+\gamma+\delta+\epsilon+\zeta+\eta$	0	1920	0	1920
θ	0	0	40	40
Totals	40	1920	40	2000

The Yulean has now reached perfect correlation, or $r=+1{\cdot}0000$.

Combine a, b, c, d, e together and f, g, h, i, j; $\alpha, \beta, \gamma, \delta$ and $\epsilon, \zeta, \eta, \theta$, and we have

(G)

	$a+b+c+d+e$	$f+g+h+i+j$	Totals
$\alpha+\beta+\gamma+\delta$	240	760	1000
$\epsilon+\zeta+\eta+\theta$	760	240	1000
Totals	1000	1000	2000

The Yulean is now *negative* and $-{\cdot}5200$.

But if we write the table as a fourfold thus:

(H)

	$a+b+c+d+e+f+g+h+i$	j	Totals
$\alpha+\beta+\gamma+\delta+\epsilon+\zeta+\eta$	1960	0	1960
θ	0	40	40
Totals ...	1960	40	2000

the Yulean would be again unity and mark perfect correlation of a *positive* kind. Mr Yule's coefficient of association would also be positive and mark perfect association.

Now it is not open to Mr Yule to turn round and assert that such tables are extremely unlikely in practical statistics, first because his condemnation of the coefficient of contingency is based solely on the creation of an artificial table in exactly the same way, and secondly because he asserts that once we dismiss the idea of Gaussian frequency the method of correlating ranks with big 'brackets' becomes applicable. Our tables bring out, however, three important points: (i) that two variates with an actual correlation of $-{\cdot}112$ may exhibit any correlation between $-{\cdot}52$ and $+1{\cdot}00$ when treated by the Yulean process of pseudo-ranks; (ii) that Mr Yule's coefficient of association may cover under the heading 'perfect association' almost any value of the real relationship, it is merely an association of common names, i.e. class-indices, and not of the real variate beneath these class-indices; and (iii) that the assumption that a fitting measure of association should give unity for a fourfold table of the form $\frac{a \mid 0}{0 \mid d}$ is by no

means obvious. It is conceivable that a better measure of association would give a limit below unity for such a case, while providing a limit nearer and nearer to unity as the information given with regard to what really occurs inside the broad categories is more and more complete. That is to say, a desirable coefficient of association would be one which would always lie numerically between 0 and 1, but which would not take the absolute value 1, unless far more detailed information were provided than is given in the statement of such a table as (H). From this standpoint we see at once how idle is Mr Yule's criticism of the coefficient of contingency. He suggests that it is invalid because (i) it has an upper limit less than unity, when the contingency table has a limited number of cells, and (ii) its value rises when you increase the number of cells. It is less than unity in the first case, because we are ignorant of what may happen when we analyse the contents of the big cells; it increases in the second case because we have additional knowledge. It only becomes unity when one character A is absolutely fixed by a second B, i.e. when A is a function of B*. The coefficient of contingency is a valid measure of association, whether the table be fourfold† or $n \times n'$-fold. It presents far fewer logical anomalies than Mr Yule's Q or the Boas-Yulean ϕ, and it readily admits of our calculating, what for many cases is essential, the probability that the two attributes are independent.

But Mr Yule dismisses the coefficient with (a) a quite unreasoned criticism that it increases in value as the number of cells increases, and (b) an illustration that it is not equal to the coefficient of correlation for one particular table of heterogeneous material, i.e. for a surface of zero correlation with a cock's comb of absolute correlation erected along its diagonal. Nobody, as far as we are aware, ever asserted it would be. The assertions made with regard to the coefficient of mean square contingency may be summed up as follows: (i) for any frequency distribution the coefficient of contingency is a reasonable measure of the extent of the deviation of the attributes from absolute independence, and (ii) for such frequency surfaces *in homogeneous material* as occur *in actual practice* the coefficient of contingency, *if the proper corrections are made*‡, gives a value close to the coefficient of correlation, whether we divide the table up into 3×3-fold or 8×8-fold groupings. The skewness of the distribution—its deviation from Gaussian frequency—is not a very disturbing factor, as we shall show in the sequel. When we take material which has—if there be an indefinite number of cells—an indefinitely great improbability of independence, i.e. material for which $C_2 = 1$, we shall not

* See Pearson, *Grammar of Science*, 3rd ed. p. 162.

† The probable error of a coefficient of contingency C_2 for a fourfold table is

$$\cdot 67449 \frac{1 - C_2^2}{\sqrt{n}} \left\{1 - 2C_2^2 + \frac{C_2(1 - \frac{1}{2}C_2^2)}{(1 - C_2^2)^{\frac{1}{2}}}\lambda\mu - \tfrac{3}{4}C_2^2(\lambda^2 + \mu^2)\right\}^{\frac{1}{2}},$$

where λ and μ have the same values as on p. 170. It does *not* become zero when the Boas-Yulean ϕ is equal to unity, unless $\lambda = \mu = 0$.

‡ These corrections have been several times referred to (see *Grammar of Science*, ed. 1911, ftn. p. 163) and have been in use in the Laboratory, but the further memoir on Contingency which has been for some time in hand has been delayed owing to pressure of other work. It will shortly be issued and deal more fully with the corrections merely stated in this paper.

reduce this infinite improbability by diluting it with a finite amount of non-correlated material; and this approach of C_2 to unity is all that Mr Yule's artificial cock's comb surface illustrates. If A is an *infinitely* improbable event, then we shall not lessen the improbability of the whole by combining the event A with B which has a zero improbability*!

Mr Yule cannot determine the efficiency of contingency methods by simply asserting that the value of the mean square contingency depends on the number of cells; it naturally alters with our increased knowledge, but this change may mark either an increase or a decrease according to the manner in which the material in the few cells is redistributed in their component cells. We assert that with a 3×3-fold table you cannot get further than the contingency, and that further progress can only be made by some other assumption as to the frequency distribution of the variates. With that assumption we think we shall be able to demonstrate in the course of this paper to the unprejudiced reader that the coefficient of contingency properly handled is, perhaps, the most powerful instrument of modern statistical theory. The assumption we make is that for correcting the results obtained by contingency, so that coefficients found for 3×3-fold, 4×4-fold, 5×5-fold, ..., 8×8-fold tables may give practically identical results, it is sufficient to deduce the required corrections by using a Gaussian hypothesis to determine certain means. The method gives excellent results for the bulk of the distributions which occur in our wide experience of statistical work. If we can show that it gives good results for the extremely skew cases which Mr Yule has gone out of his way to cite, our point will be proved. Since the full development of the contingency method, fourfold tables have not been used by the Biometric School, except as controls, where contingency tables could be formed on the given data. But the statement that contingency was developed in order to overcome the difficulties of the fourfold table methods is directly disregarded by Mr Yule when he turns to our pigmenta-

* A is the probability that in examining two absolutely independent variates, n cells shall be occupied and $n^2 - n$ cells empty when we make n infinite. B is the probability that the n^2 cells shall each have their theoretically independent contents. No combination of these two events can give less than an indefinitely great improbability, i.e. $C_2 = 1$. But we anticipate that if Mr Yule had not raised his cock's comb at such a conspicuous angle to the rest of his surface that its heterogeneity would be readily visible to the trained statistician, there would be no very serious error introduced by applying the mean square coefficient of contingency even to moderately heterogeneous material. We have not had leisure to investigate the matter closely, but if we superpose two Gaussian frequency-surfaces with identical means, with the same standard deviations for both variables, and with correlations r_1 and r_2, then the true correlation by product moment is

$$\rho = pr_1 + qr_2,$$

where pN and qN are the total frequencies of the two components and $p + q = 1$.

On the other hand

$$C_2 = \sqrt{\frac{\rho^2 (1 + r_1 r_2) - 2\rho r_1 r_2 (r_1 + r_2) + r_1^2 r_2^2 (1 + r_1 r_2)}{\rho^2 (1 + r_1 r_2) - 2\rho r_1 r_2 (r_1 + r_2) + r_1^2 r_2^2 (1 + r_1 r_2) + (1 - r_1^2)(1 - r_2^2)(1 - r_1 r_2)}}.$$

If $r_1 = 1$, $r_2 = 0$, as we know $C_2 = 1$, while $\rho = p$. But if $r_1 = \cdot 2$, $r_2 = \cdot 7$, with $p = \cdot 3$, $q = \cdot 7$, the mixture proportions of Mr Yule's illustration, then $\rho = \cdot 55$ and $C_2 = \cdot 59$. If $r_1 = \cdot 5$, $r_2 = \cdot 7$, $p = \cdot 3$, $q = \cdot 7$, then $\rho = \cdot 64$ and $C_2 = \cdot 65$. Again if $r_1 = \cdot 7$, $r_2 = \cdot 3$, $p = \cdot 4$, $q = \cdot 6$, then $\rho = \cdot 46$ and $C_2 = \cdot 49$. Thus it does not appear that small amounts of heterogeneity, *not* detectable on a study of the table, are likely to give very misleading values when C_2 is taken as a practical measure of ρ.

tion data*. The recognition, however, that the fourfold table may give discordant results—a recognition made by the Biometric School within four years of the publication of the pigmentation investigations for eye-colour in man and coat-colour in horses—does not dismiss the fourfold table from practical statistics, but only from that portion of it where multiple contingency tables are available. Given a fourfold classification alone, how is it to be treated? We reply unhesitatingly that in the great bulk of cases the use of tetrachoric r_t is the best treatment. We base this on the experience that where nothing is known the Gaussian is far more likely to describe approximately the frequency than any other hypothesis. Even taken as a mere coefficient of association, tetrachoric r_t is better than Mr Yule's Q or the Boas-Yulean ϕ, except for absolutely discrete units as in the purely theoretical Mendelian cases; and in those cases the correlation of ranks is the correlation of variates, as Pearson indicated in his memoir of 1904†.

(5) *On the Surface of Constant Association and on "Natural" Equalisation.*

As we have indicated, Mr Yule never states adequately the conditions under which his coefficients of association and colligation are to be applied. He apparently considers the nature of the continuity of his frequency surface, if his variates are continuous, to be absolutely immaterial. Now in the case of every two continuous variates, whatever their nature, a frequency surface does exist for which Mr Yule's association or colligation is constant wherever the divisions may be taken upon which the fourfold table is based. Let n_{pq} be all the first quadrant frequency corresponding to the total frequencies p and q of the two variates, where p and q are supposed to be absolutely known. If the association coefficient be Q and the notation

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	N

$$a+b=q,$$
$$a+c=p,$$

we have $$(1+Q)/(1-Q)=ad/bc=\chi, \text{ say.}$$

* In the very same number of *Biometrika* (Vol. III. 1904) in which the Huxley Lecture appeared, there is a paper on the inheritance of pigmentation in the Greyhound; it is the work of Pearson's Laboratory and started about the same time as the Huxley Lecture reductions. The following words occur: "When we first started work on the greyhounds, the method of contingency had not been developed, and accordingly we made tables for the inheritance of melanism and of red pigment and proceeded to find the correlations by the fourfold division process" (*l. c.* p. 252). And again "In order to compare the fourfold method with contingency methods, 16-fold tables and 25-fold tables were worked out to compare with the fourfold tables adopted for the inheritance of red and black pigment respectively" (p. 253). "The results deduced by contingency D method are singularly uniform and steady as compared with those of the fourfold-table methods, and we believe, if it be adopted generally for such pigmentation problems, it will not only free us from any question of pigmentation scale, but afford a good result on a not excessive expenditure of calculating energy" (p. 253). It is clear that the Laboratory publicly admitted the difficulties of the fourfold-table method *two years* before Mr Yule started to criticise it as applied to pigmentation data! Yet Mr Yule never mentions this fact.

† "On a Generalised Theory of Alternative Inheritance with special reference to Mendel's Laws," *Phil. Trans.* Vol. 203, p. 53.

Then $\quad a = n_{pq}, \quad b = q - n_{pq}, \quad c = p - n_{pq}, \quad d = N - p - q + n_{pq}.$

Therefore $\quad (\chi - 1)\, n^2_{pq} - n_{pq}\, \{(\chi - 1)(p + q) + N\} + pq\chi = 0$

is a quadratic to find n_{pq}. If we take as we always can arrange to do Q *positive*, then χ lies between 1 and ∞ or is also positive. The equation for n_{pq} will have real roots if

$$\{(\chi - 1)(p + q) + N\}^2 > 4\chi(\chi - 1)pq,$$

or, if

$$(\chi - 1)^2 (p - q)^2 + (\chi - 1)(p - q)^2 + (2N - p - q)(\chi - 1)(p + q) + N^2 > 0,$$

which is true since $\chi > 1$ and $2N > p + q$. Hence, since $\chi pq/(\chi - 1)$ is always positive, the quadratic has always one and only one real positive root.

Further: $$\frac{\delta n_{pq}}{dp} = \frac{\chi q - n_{pq}(\chi - 1)}{N + (p + q - 2n_{pq})(\chi - 1)},$$

but χ is $> \chi - 1$, $q > n_{pq}$, and $p > n_{pq}$, therefore it follows that both numerator and denominator are positive, or δn_{pq} increases with p. Similarly it increases with q, or in subtracting n_{pq} from either $n_{p+\delta p,\, q}$ or $n_{p,\, q+\delta q}$ we shall never reach a *negative* difference. Thus it is always possible to construct a surface for which Q is constant for every fourfold division. It seems to us that had Mr Yule realised the possibility of this surface and studied it*, he would have known more about the real properties of his Q and its bearing on such distributions as occur in practice. The fact that in every distribution of continuous variates we come across there is no approach to constancy in Q, that it varies continuously and almost in a predictable manner shows how very far the surface of constant Q is from representing the facts of experience. Still had Mr Yule fitted the best surface of constant Q to a known distribution of detailed data, and so ascertained his value of Q, he would have given us a coefficient which would have lived in statistical practice and theory, and he would have thrown real light on the relationship of association to correlation. We have not spent time in discussing the complicated equation to the surface of constant Q, but we have provided one illustration of such a surface. Taking the total frequencies of each eye-colour group in Father and Son, only adding 5 and 6 together, we have Table I given below for the eye-colour distribution categories; this table would within the limits of our decimal places have the same coefficient of association, $Q = 0{\cdot}6$, wherever we divide it into fourfold tables. For example, taking both divisions between 2 and 3 we have the fourfold

191·55	143·45	335
166·45	498·55	665
358	642	1000

$Q = {\cdot}59996.$

* We suggest that Mr Yule has not studied the matter, for he writes: "There is one case and one only where Q is independent of the axes chosen, and that is where the variables are strictly independent," *Phil. Trans.* Vol. 194, p. 278. For the equation to the surface see Appendix III.

Or, again, between 6 and 7 vertically and 3 and 4 horizontally

576·60	42·40	619
294·40	86·60	381
871	129	1000

$Q = \cdot 60002$,

which sufficiently illustrate the feasibility of the construction of the surface.

But at once the fictitious character of Mr Yule's idea of association comes to light. He tells us that it is "natural" to put a table in the equivalent symmetrical form—one for example in which we have 50 °/₀ deaths and 50 °/₀ recoveries, 50 °/₀ vaccinated and 50 °/₀ unvaccinated. And this is possible because his Q does not change by such a process. But surely if it is "natural" to have equal numbers of deaths and recoveries, it is also "natural" to have equal numbers of fathers, and for the matter of that equal numbers of sons, in each eye-class. Equal light blue eyes, and equal dark blue eyes, equal light brown eyes, equal dark browns and blacks. But the instant we study a table of this kind some

TABLE I.

Table of Variates in Father and Son Eye-Classes for Constant Association Coefficient 0·6.

	1	*2*	*3*	*4*	*5+6*	*7*	*8*	Totals
1	4·08	19·04	6·25	2·60	0·78	1·00	0·25	34
2	19·54	148·89	75·33	32·54	9·68	12·05	2·97	301
3	7·41	90·46	90·17	52·97	16·86	21·03	5·10	284
4	2·19	28·54	39·43	33·19	12·53	16·89	4·23	137
5+6	1·35	17·12	25·63	26·40	11·82	17·83	4·85	105
7	1·03	13·15	19·94	23·33	12·18	21·60	6·77	98
8	0·40	4·80	7·25	8·97	5·15	10·60	3·83	41
Totals	36	322	264	180	69	101	28	1000

remarkable points arise. (i) The selection which is advocated by Mr Yule and which keeps his coefficient constant is a perfectly arbitrary one, and (ii) it must be stopped at a perfectly artificial limit, namely that at which the arbitrary division has been made. Mr Yule does not equalise *each* sub-group, but if, for example, he divides into light and dark eyes between 3 and 4, he multiplies his light blue eyes, his dark blue eyes and his greys, not by different factors which would equalise these groups *with each other*, but by the same common factor, so that the sub-groups really remain in the same proportions as before the equalising. Surely if it is "natural" to equalise light and dark eye totals, it is equally "natural" to equalise the totals of light blue and dark blue eyes. If it is "natural" to equalise the vaccinated and unvaccinated totals, it is surely for the same reason "natural" to equalise the several groups of individuals who have one, two, three vaccination scars, for the number of their scars "is dependent to a large degree on a purely arbitrary circumstance, the activity of the authorities" in making 1, 2, or 3 punctures. It is equally "natural" to equalise the numbers who have lived 5, 10, 20, 30

or 40 years since their last vaccination, for this "is dependent to a large degree on a purely arbitrary circumstance," the question of whether the population has been recently alarmed and revaccinated. The efficiency of vaccination depends both on size of cicatrix and on interval since vaccination. Yet if Mr Yule makes these equalisations he will wholly change his association coefficient. That coefficient admits not of a natural selection, but only of a wholly artificial one, namely, one which alters all sub-classes on one side of an arbitrary dividing line in the same ratio and not in their "natural" ratios. Every word Mr Yule has uttered with regard to equalisation and its "naturalness" applies equally to classifications in 3, 4 or more groups. Why should the relationship of father to son in eye-colour depend upon the number of light blue eyed fathers taken? It is "natural" to equalise them with the dark blue. But this is only possible if the multiplying of a row or column will not influence the result, and the reader has only to test on the above table how Q is changed when he starts such a selection. A few of the values we have obtained are $Q = \cdot47, \cdot44, \cdot36, \cdot17$ and $\cdot04$! The fact is that Mr Yule's is not a *general* selection, but a perfectly artificial one, which is not in the least "natural" when we analyse its effect on the constituents of a given class.

It is as well to stop and inquire what influence this "natural" equalising has on contingency tables. Below (Table II) is given the table for eye-colour in Father and Son reduced to five classes on account of the labour involved. The equalising factors are

For the Rows:	*For the Columns:*
$y_{1+2} = 3{\cdot}73512$,	$x_{1+2} = \cdot61949$,
$y_3 = 3{\cdot}92073$,	$x_3 = \cdot81174$,
$y_4 = 7{\cdot}11764$,	$x_4 = 1{\cdot}00000$,
$y_{5+6} = 7{\cdot}47842$,	$x_{5+6} = 2{\cdot}38871$,
$y_{7+8} = 6{\cdot}35625$,	$x_{7+8} = 1{\cdot}34437$,

and the resulting table is given as Table III.

TABLE II.

Actual Eye-Colour Table for Father and Son in Five Classes.

Eye-Colour of Father*.

Eye-Colour of Son.	*1+2*	*3*	*4*	*5+6*	*7+8*	Totals
1+2	194	70	41	9	21	335
3	83	124	41	13	23	284
4	25	34	55	11	12	137
5+6	27	12	19	24	23	105
7+8	29	24	24	12	50	139
Totals	358	264	180	69	129	1000

* See *Phil. Trans.* Vol. 195 A, p. 138.

TABLE III.

Table for Eye-Colour in Father and Son equalised, or put into Mr Yule's "Natural" Groups.

Eye-Colour of Father.

Eye-Colour of Son.	*1+2*	*3*	*4*	*5+6*	*7+8*	Totals
1+2	448·88	212·23	153·14	80·30	105·45	1000
3	201·60	394·65	160·75	121·76	121·24	1000
4	110·24	196·44	391·47	187·02	114·83	1000
5+6	125·09	72·85	142·09	428·73	231·24	1000
7+8	114·19	123·83	152·55	182·19	427·24	1000
Totals	1000	1000	1000	1000	1000	5000

In this case the Q was not constant for all divisions at starting, and therefore there is no general standard. But Mr Yule tells us that his method of pseudo-ranks is the best known to him for such a table. The Yulean for this table before equalisation was ·403, after equalisation it is ·166! Q can change as much as from ·52 to ·43. It is only an "unnatural" equalisation which keeps Q constant.

Let us consider another case of this, taken from the one field to which Mr Yule in his present paper ventures to apply his coefficient, that of vaccination*.

TABLE IV.

Severity of Small Pox.

		Haemorrhagic	Confluent	Abundant	Sparse	Very Sparse	Totals
Vaccinated. Years since vaccination	*0—10*	0	1	6	11	12	30
	10—25	5	37	114	165	136	457
	25—45	29	155	299	268	181	932
	Over *45*	11	35	48	33	28	155
Unvaccinated	...	4	61	41	7	2	115
Totals ...	...	49	289	508	484	359	1689

The equalisation factors are:

For Rows:	*For Columns:*
$y_1 = 6{\cdot}1300$,	$x_1 = 11{\cdot}0816$,
$y_2 = {\cdot}3631$,	$x_2 = 1{\cdot}5125$,
$y_3 = {\cdot}1496$,	$x_3 = 1{\cdot}0000$,
$y_4 = {\cdot}6923$,	$x_4 = 1{\cdot}0113$,
$y_5 = 1{\cdot}0697$,	$x_5 = 1{\cdot}1659$,

reckoning rows and columns from the left-hand top corner.

* See *Biometrika*, Vol. VII. p. 257.

After equalising the classes of this table we have the following one:

TABLE V.

Vaccination Grade and Intensity of Disease.

Severity of Small Pox.

		Haemorrhagic	Confluent	Abundant	Sparse	Very Sparse	Totals
Vaccinated. Years since vaccination	*0—10*	0	9·27	36·78	68·19	85·76	200
	10—25	20·12	20·32	41·39	60·59	57·58	200
	25—45	48·08	35·07	44·74	40·54	31·57	200
	Over *45*	84·40	36·66	33·23	23·11	22·60	200
Unvaccinated	...	47·40	98·68	43·86	7·57	2·49	200
Totals ...	...	200	200	200	200	200	1000

Now here are some of the changes that take place in Q, owing to this equalisation of the classes:

TABLE VI.

Vertical Division	Horizontal Division	Old Q	New Q
Confluent-Abundant ...	Vaccinated-Unvaccinated	− ·7220	− ·7070
Abundant-Sparse	Vaccinated-Unvaccinated	− ·8599	− ·8945
Confluent-Abundant ...	25—45 and over 45 ...	− ·5714	− ·7522
Confluent-Abundant ...	10—25 and 25—45 ...	− ·5411	− ·8163
Abundant-Sparse	10—25 and 25—45 ...	− ·4469	− ·7742
Haemorrhagic-Confluent...	10—25 and 25—45 ...	− ·5711	− ·7798
Haemorrhagic-Confluent...	25—45 and over 45 ...	− ·5449	− ·5861
Haemorrhagic-Confluent...	Vaccinated-Unvaccinated	− ·1009	− ·1371

The range of values found for Q for a single table will be considered later; it is sufficient here to indicate how largely the bulk of them are changed by equalisation of the groupings. If further evidence be needed for the radical changes which invariably accompany Mr Yule's conception of a "natural" grouping, we may note that the Yulean deduced from pseudo-ranks is before equalisation − ·3099, but after equalisation − ·5375. These two equalisations of class frequency show that it is impossible to predict *a priori* how the relationship of the two variates will be changed by the process; in the first the relationship was lowered by 59 °/。 and in the second case raised by 73 °/。 of its value as estimated by the Yulean*! It is clear that only a fictitious type of equalisation has been used by Mr Yule. His independence of selection is only an artificial one; sub-groups within his categories are not equalised but only equally selected; and further such selection must neither fall short of nor exceed the arbitrary division he has selected for his class. If for example we agree to equalise each

* Consider as a last illustration the Table VII below. The correlation is ·37. The reader will find

age group of the vaccinated—surely a "natural" process—then we shall modify the coefficient of association obtained by dividing between vaccinated and unvaccinated. If we equalised the totals of vaccinated and unvaccinated, then Q would have changed, if the division were taken at an interval of 25 years since vaccination. The whole process is quite arbitrary, and we believe wholly without validity.

(6) *On the Coefficient of Association and the Assumption of Discrete Variates.*

Mr Yule has asserted that when we free ourselves from any necessary relation to the theory of normal correlation, the ordinary theory of correlation is applicable in its entirety to the 2×2-fold table. Apparently he holds that the same is true for his coefficients of association and colligation although these different methods lead to diverse and often contradictory results, for the simple reason that selection vastly affects one and does not affect the other. He has not considered the surface of equal association, nor discussed the cases to which the "four-point" surface, involved in using ϕ as Boas and he propose, can be legitimately applied. It is accordingly of interest to see what happens when the coefficient of association is applied to various types of frequency surface.

It is no valid reply to criticisms based upon such an investigation to say that such frequency surfaces do not occur in practice. Mr Yule has never entered into any discussion of the character of the distributions to which he applies his association; it is sufficient for him that they give a fourfold table, and he makes no appeal, as we do, to experience as a basis for the adoption of any coefficient. It is therefore possible to test his association against the clear idea of correlation on any distribution whatever.

If Mr Yule replies that this is not fair treatment because the coefficient of association applies to discrete quantities, we answer that he has never defined

it an amusing task to equalise the total frequencies; he will then discover that it takes an interesting form, that of an old friend of Mr Yule's, and the correlation will then be recognised as ·50.

TABLE VII.

First Variate.

Second Variate.		*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	Totals
	a	11	20	30	35	40	45	50	40	30	2	303
	b	20	4400	600	700	800	900	1000	800	600	40	9860
	c	30	600	9900	1050	1200	1350	1500	1200	900	60	17790
	d	35	700	1050	13475	1400	1575	1750	1400	1050	70	22505
	e	40	800	1200	1400	17600	1800	2000	1600	1200	80	27720
	f	45	900	1350	1575	1800	22275	2250	1800	1350	90	33435
	g	50	1000	1500	1750	2000	2250	27500	2000	1500	100	39650
	h	40	800	1200	1400	1600	1800	2000	17600	1200	80	27720
	i	30	600	900	1050	1200	1350	1500	1200	9900	60	17790
	j	2	40	60	70	80	90	100	80	60	44	626
	Totals	303	9860	17790	22505	27720	33435	39650	27720	17790	626	197399

discrete quantities, and we can only judge what he means by this term from seeing the cases to which he has applied it. We shall therefore deal first with the discussion of the cases to which Mr Yule has applied his methods, and then consider their effects as applied to (*a*) continuous, (*b*) discrete variates.

Mr Yule opens his paper by observing* that if we classify objects into two classes only, for example "peas into yellow-seeded and green-seeded, or the members of any group of mankind into male and female, the resulting data are of the simplest possible form." The data may be thus in the simplest possible form, but difficulties might occur even in such simple cases as those cited by Mr Yule. The classification even into yellow- and green-seeded peas is by no means so simple as Mr Yule suggests, and certain types of hermaphrodite in man might undoubtedly puzzle even Mr Yule's powers of discrimination. To understand really what Mr Yule proposes to classify we must turn to the cases to which he has applied his method. They are as far as we can judge from an examination of his writings the following: (1) Good and Bad Temper, (2) Presence and Absence of the Artistic Faculty, (3) Stature in Man, (4) Tallness in Plants, (5) Mental Dulness, (6) Low Nutrition, (7) Defects in Development involving "size, form or proportioning of parts," (8) Abnormal Nerve Signs, involving "abnormal actions, movements, and balances," (9) Mental Derangement involving imbecility and idiocy, (10) "Blindness," (11) Deaf-mutism, (12) Recovery and Death in the case of Smallpox, (13) Vaccination or non-vaccination, (14) Male and Female, (15) Cross and Self-Fertilisation, (16) Eye-Colour, (17) Colour of Flower, (18) Prickliness of Fruit. Mr Yule has probably used or suggested the use of the coefficient of association in other cases†. Looking through the above cases we see it is in the rarest instances, possibly only in (14) and (15), that Mr Yule has confined himself to discrete variates. He has applied his coefficient of association over and over again to continuously varying quantities. Temper (1), artistic faculty (2), stature (3), tallness (4) are all quantitatively determinable variates, even if difficult in some cases to measure. One man has a better or worse temper than another; one man has a greater or less degree of artistic faculty; where the division between presence and absence of the faculty is put, or what is called good or bad temper is largely matter of personal equation. But nobody doubts the range of such variates; they are continuous, there is no sudden break.

Now turn to the next four characters (5)—(8). No one who has studied the essential difficulty of defining what is feeble-minded will doubt the continuity of mental dulness. It is not a discrete character, but a continuous variate. There are certainly all grades of mental defect, and the groups idiot, imbecile, feeble-minded, "simple" are quite artificial. If the whole population were graded according to intelligence, the frequency curve would be continuous; no one knows whether it would be Gaussian or not, or whether it would be "humpy" towards

* *Journal R. S. S.*, Vol. LXXV. p. 579.

† He has apparently approved of its application, as we shall see later, to a long series of absolutely continuous variates by Professor Niceforo.

the tails. If there is a gap between the defective and the normal members of the same sibship, intermediates in plenty will be found elsewhere. The whole difficulty associated with the Government Bill for the care of the feeble-minded turns on the questions how and where to draw the line between the "normal" and "feeble-minded." Whatever may be finally done, it is quite certain that there will be no real distinction between two individuals who fall just one side and just the other of the dividing line. Even the personal equation of trained observers will vary in classification, and Mr Yule takes the untrained record of thousands of the laymen who make census returns as marking off in some manner a "discrete" character of "mental defect." Absolutely the same remark applies to nutrition. The boundary of "low nutrition" is also an arbitrary dichotomy in a continuous variate fixed by the personal equation of the observer. (7) tells its own tale, for it is based on continuously changing and measurable characters. (8) is less obvious, but not only in number, but in quantitative intensity "nerve signs" are really continuous. To these also we may add mental derangement (9); there is a very great range of variation in imbecility and idiocy as we have already indicated under mental dulness. If we turn to "blindness" (10) the source of it may be most varied, but if we define it merely as the loss of the faculty of sight, there is, as the simple cases of either congenital or senile cataract might have shown Mr Yule, almost every intensity of the failure of sight. Even certain cases of albinism are to be found in the Blind Asylums and are educated as semi-blind. Semi-blindness is so well recognised that special schools have been started for the semi-blind by certain educational authorities. In Bristol out of 22 children sent to the Blind Asylum at the expense of the Education Committee 13 had some degree of vision, and 7 could read large print with varying degrees of difficulty, and were able to do some form of handwork by sight. Out of 75 children specially examined for eye defects at Bristol in 1911, beyond those requiring glasses were 9 suitable for the partially blind class*. Of (11) deaf-mutism, we are less competent to speak, but we have been informed by the very best authorities it is far from an absolutely fixed condition and that it has a great variety of grades. The grades are more marked in the acquired than the congenital cases, but in the latter cases they vary with the cause of the congenital deafness. Cases even exist with slight degrees of deafness which would probably have been associated with mutism had the deafness been more considerable. Different degrees of hearing are found among deaf-mutes. There are scores among them who undoubtedly possess an amount of vowel hearing, and it helps the tone of their voices when they are being taught articulation; this is true, although they cannot distinguish speech without watching the lips and without their ears being within two or three feet of their interlocutor.

* We have had before us the diary of a man who "went blind" in old age. At 70 the writing is perfectly clear and legible. It closes at 80 with the words scrawled in an almost unreadable trembling hand across the page. "I am now almost blind and with great difficulty I write this. Oh! the misery I feel, no one comprehends it.—" When was this man "blind" for a census return? The category "blind" covers a vast range of cases of partial sight, and it would be hard to draw a dividing line between "blind" and "normal" in such cases.

The number unable to articulate is negligible; all the deaf can make sounds and probably 90 °/₀ could articulate with more or less success. Those who acquire deafness later in life retain their speech, but with impaired quality. The census returns club together the congenital and acquired forms. Where then is the "discrete" attribute? As for the blind so for the instruction of the semi-deaf, special schools have been established by some educational authorities. There were 19 scholars reported on in the Bristol Education Committee's Report for 1911 as attending the semi-deaf school at Broad Weir. From "very deaf" to "very slight deafness" we have every graduation of hearing from those that can only hear under 1 yard to those who can hear at 6 yards. There is every variation in speech from the "unintelligible," "now beginning to use a little," "voice very weak," up to "lisping nearly overcome" and "speech of good quality." Speech-reading forms an essential feature of the instruction, and the cases are those of transitional deaf-mutism.

We entirely disagree with Mr Yule's statements that such attributes as blind and seeing, deaf and not deaf, mentally deranged and not deranged, "if not absolutely discrete, are very largely distinct from each other" (*loc. cit.* p. 638), and we wholly fail to follow his argument on this point.

Mr Yule has apparently seized on recovery and death in cases of small-pox as discrete instances, but by Dr Macdonell and one of the present writers they were used to measure a continuous quantity—the severity of the attack. The data due to Dr Brownlee, and published by one of us in *Biometrika*, Vol. VII. p. 256, show that when the severity of the attack is classed by such categories as haemorrhagic, confluent, abundant, sparse, very sparse, that variate is essentially continuous, and that the mortality is largely confined to the two highest classes. Again when the immunity conferred by vaccination is reclassed under "unvaccinated," vaccinated over 45 years ago, 25—45 years back, 10—25 and 0—10, we at once recognise that vaccination regarded as conferring immunity is an essentially continuous variate. The same notion of continuity comes in, if we classify severity by the number of pustules on the face, comparison being made with a standard series of photographs of typical cases, or again show that area of vaccination scar affects the extent of the immunity. Mr Yule can hardly be ignorant of all this work, yet he lightly chooses the vaccination data as good illustrations of his method*. The classifications of intensity of attack by number of pustules and of vaccination by period since vaccination show in general frequency surfaces of a rough "cocked hat" type, and dispel at any rate the notion that immunity and severity can be treated as discrete variables. It is only the confirmed Mendelian who would classify any pigmentation character, whether of eye or of coat or of flower, into two alternative groups. The intensity of pigment is undoubtedly a continuous

* That he was, as long ago as 1899, acquainted with the varying intensity of vaccination is conclusively proved by the following sentence from his memoir on 'Association' (*Phil. Trans.* Vol. 194 A, p. 289): "The association between non-vaccination and attack is very high indeed for young children —·8 to ·9—but drops sharply to ·5 (*owing presumably to the waning protection of the vaccination made in infancy*) in the older age groups." The italics are ours.

variate, when we consider the actual frequency of pigment granules, and it is only confusing the issue when eyes are divided into two classes, those with both posterior and anterior pigment, and those with only posterior pigment. We shall return to the question of eye-colour later, as it is a case out of which Mr Yule makes much capital, but which in his last paper is the only one he treats by his own association methods. If we take "prickliness" in fruit, there is no evidence yet that it has ever been properly measured, and that it would not prove to be a continuous variate, much as "hairiness" proved to be in the case of *Lychnis* when actually measured by Weldon*. To sum up, there are not among the attributes used by Mr Yule to illustrate association coefficients any but those of sex and nature of fertilisation which can reasonably be considered discrete quantities, and in using even these he always couples them with characters which in our opinion are distinctly continuous variates.

We shall therefore start this criticism of Mr Yule's statistical investigations by indicating the fallacious nature of his coefficients of association and colligation as applied to continuous variates. We shall then deal with the question of discrete alternative variables, and show their absurdities in that case. Finally we shall show reason for questioning the details of the bulk of his memoir, which is not occupied with the discussion of his special coefficients at all, but in advocating a new empirical method which there is ample reason for considering equally fallacious.

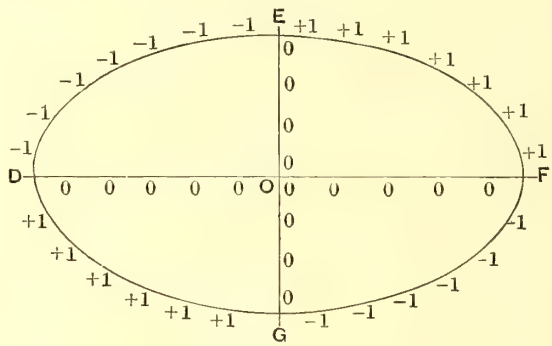
(7) *On the Idleness of Mr Yule's Coefficient of Association when applied to Continuous Variates.*

(*a*) *The Need for either Knowledge or Hypothesis as to the Nature of the Frequency.*

Let us start from a very general case of two continuous variates. The first question that we require to answer is whether for a given value of one variate the mean value of the other variate changes. If we can find the mean values of the arrays of the second variate for given values of the first we obtain the regression line; should this be a straight line the correlation coefficient r is a suitable measure of the relationship, if it be not then the correlation ratio η gives us a measure, and $\eta^2 - r^2$ marks the deviation of the regression from linearity. In cases where the regression lines are both straight and $r = 0$, it by no means follows that the two variates are absolutely independent. The next essential condition is the equal variability of the arrays, or what is nearly the same thing the probability of a combination of one variate x between x and $x + \delta x$ with the other variate y lying between y and $y + \delta y$ is the product of the probabilities

* *Biometrika*, Vol. II. pp. 47—55. The danger of this sort of classification has come home very emphatically to one of the present writers, who had tried to use the category of "short-muzzle" as against "long-muzzle" in breeding dogs. He was convinced that "short-muzzle" was a dominant character in the Mendelian sense, until he took actually to measuring the muzzles of the hybrids of first and later generations, when the idleness of treating such categories as discrete quantities became at once obvious.

of x lying between x and $x + \delta x$ and of y lying between y and $y + \delta y$. This is the true measure of independence of the two variates. It is always possible to test the independence of two variates—i.e. the probability of their independence—by aid of the mean square contingency and the use of Palin Elderton's Tables*. But there is a whole class of continuous variates for which the regressions are linear and the correlation coefficient is zero, but which are yet heteroscedastic, i.e. the arrays of the y-variate for a given x-variate are not similar frequency distributions. Any frequency surface, with two planes of symmetry, one perpendicular to the axis of measurement of each variate, is representative of this class. Let $DEFG$ be the oval contour line within which, for a given population N, all the frequency lies. This may be the actual curve in which the frequency surface cuts the plane of xy, or if the surface asymptotes to that plane, the contour within which all individuals of the N observed actually lie. If the distribution were Gaussian, this contour would be an ellipse. Generally let us suppose it any non-re-entering oval curve†.



Let the frequencies in the four classes made by divisions parallel to the axes of the variates, i.e. to DOF and EOG, be represented by $\dfrac{a \mid b}{c \mid d}$, which we have used throughout in preference to Mr Yule's more cumbrous $\dfrac{AB \mid A\beta}{\alpha B \mid \alpha\beta}$.

Then Mr Yule's coefficient of association Q is $(ad - bc)/(ad + bc)$; it is unnecessary at this point to consider Mr Yule's coefficient of colligation ω, which is only a special function of Q.

All along EOG taken as one dividing line (the other will be perpendicular to it), $Q = 0$; all along DOF also, Q is zero. All along the arcs EF and DG, $Q = +1$; all along the arcs FG and DE, $Q = -1$. These values are indicated by numbers in the diagram. If the frequency surface be non-re-entering, then when the axis of the rectangular dividing planes is taken anywhere in the quadrants EOF and DOG, Mr Yule's Q is *positive* and varies from 0 to 1; if this axis be

* Pearson, *Drapers' Research Memoirs*, "On the Theory of Contingency," p. 6 (Dulau and Co.), and *Biometrika*, Vol. I. p. 155.

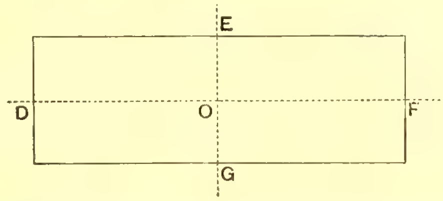
† If the curve be a re-entering one, the rapid value changes of Mr Yule's coefficients are still more remarkable!

taken in the quadrants *EOD* and *FOG*, Q is *negative* and varies from 0 to -1. Hence for two continuous variates, whose actual correlation is zero but which are not homoscedastic, we find Mr Yule's coefficient will run through the whole range of possible values from -1 to $+1$, and what its observed value will be depends solely on where we take our dividing planes.

It is perfectly true that Mr Yule's Q vanishes if the variables are *theoretically* independent, but no variables are *practically* independent, and in actual statistics of two continuous variates when the grouping is fairly small we do not get *DEFG* a rectangle, the *sole* bounding contour for which Mr Yule's coefficient of association is zero *all round*. The fact that Mr Yule gets $+1$ or -1 for his coefficient round his boundary contour would be of small importance were it not that he appears to hold that, when $Q = \pm 1$, then its probable error is zero. Round the bounding contour of a distribution of this kind Pearson's normal coefficient has usually a big probable error, and the investigator is thus warned that its vagaries are of no account*. When the investigator comes, however, to a fourfold classification leading to $Q = \pm 1$ by the zero of one of its classes, he would, if he were to follow Mr Yule, assume his result absolutely reliable, and to be so independently of the total population used. As a matter of fact, his dividing lines may have given $Q = \pm 1$ solely because they chanced to be taken near the bounding contour of a frequency distribution, of which the investigator knows nothing! The real fact of the case is that Mr Yule's investigation of the probable error of Q, while correct as long as the frequency of any of his four classes is substantial, fails entirely when one of his four classes is zero, and is correspondingly in error when Q is very large owing to one of the four classes being very small. Even if Mr Yule determines the probable error of Q for such cases by higher approximations, it will be meaningless without consideration of the frequency distribution of Q, which is an exceedingly skew curve for high numerical values of Q.

Before we leave these cases of zero correlation it is worth while to indicate how Q works for various artificial frequency surfaces.

(1) A rectangular block. Q is zero all round the boundary and for all possible divisions.

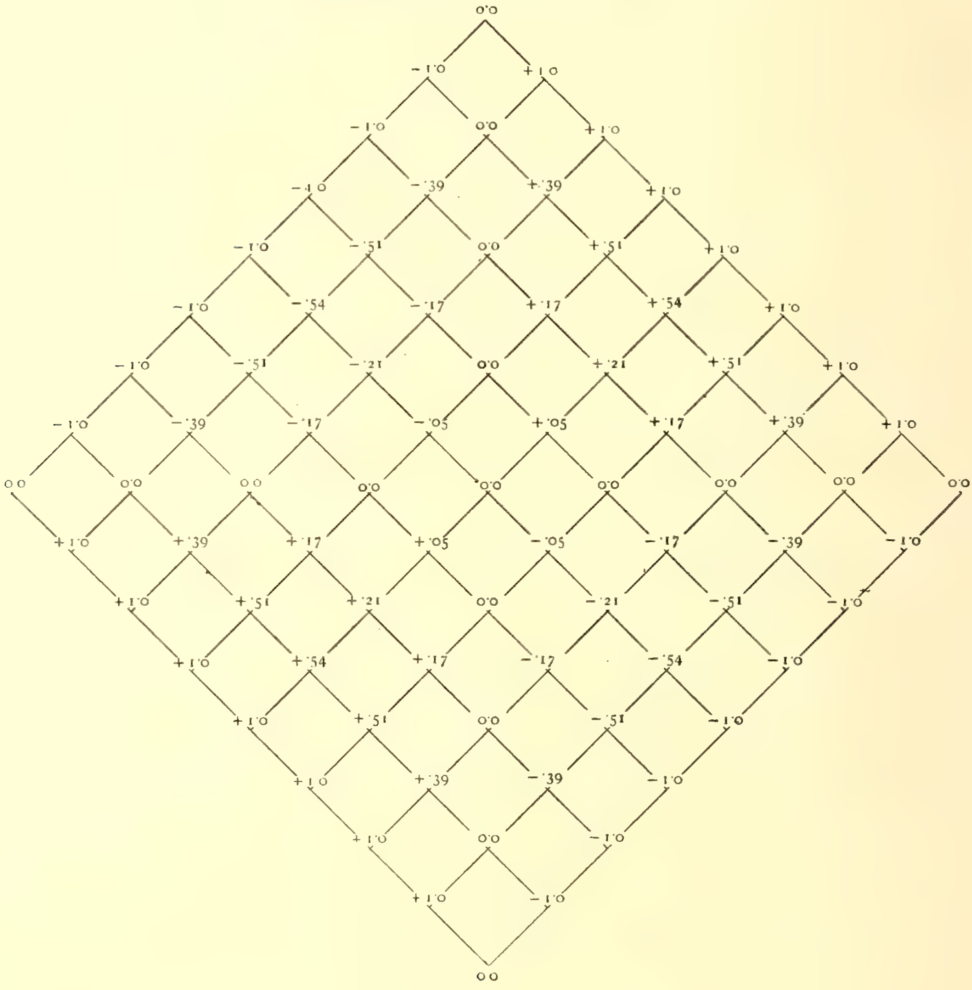


(2) A square prism; diagonal planes parallel to the variate axes. Q is positive in two quadrants and negative in two, and takes all values between $+1$ and -1 according to the point of division. It is essential to note that Q does not, as

* The matter is discussed more at length in Appendix I.

might be thought, give small values except in immediate proximity to the boundaries, it rises to quite substantial values when the percentage in any quadrant takes a value which Mr Yule has not hesitated to use when criticising other methods of approaching the problem of association, and—according to Mr Yule—as it approaches these values the reliability of his coefficient increases!

DIAGRAM II. Frequency surface of zero correlation exhibiting every possible variation of Q with different dichotomic lines.



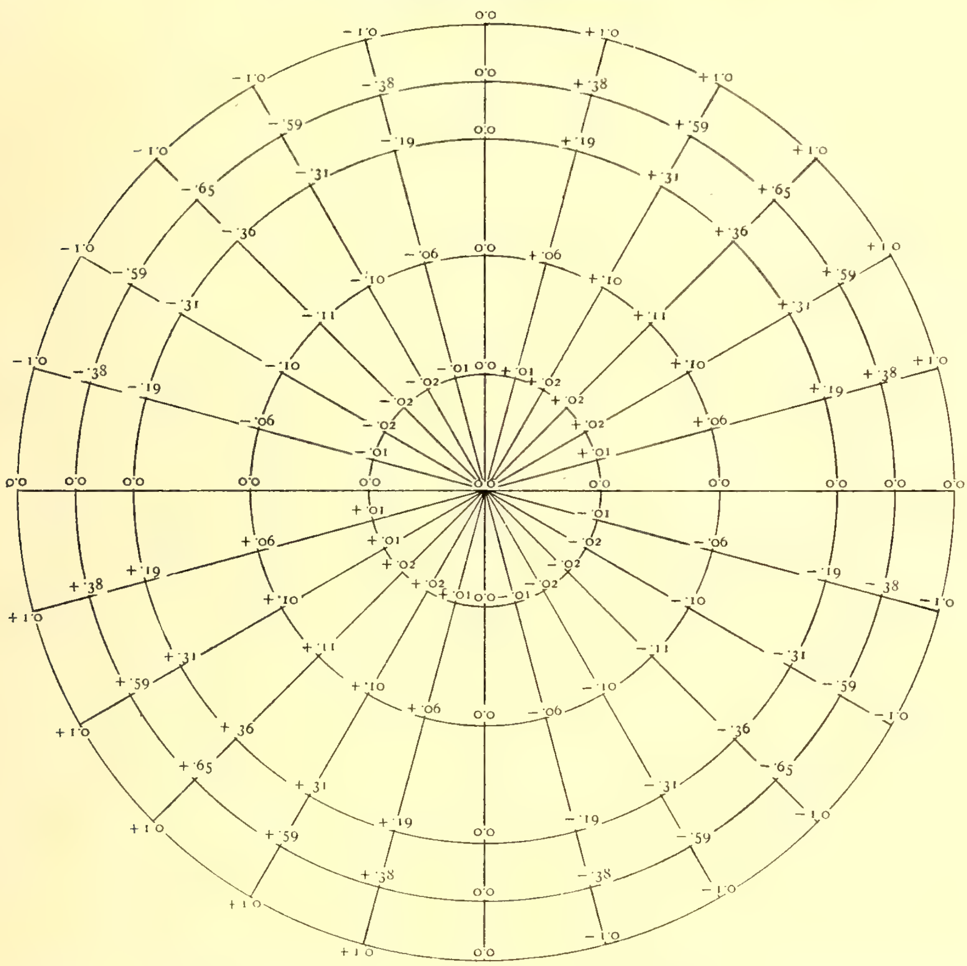
(3) Let us now take a cylinder on circular base. Here again the correlation coefficient is zero, but Mr Yule's coefficient runs through its whole range from -1 to $+1$, being negative in two quadrants and positive in the other pair.

The above cases illustrate, we think effectively, the point that the coefficient of association tells us absolutely nothing, unless we make some assumption as to the

nature of the surface of frequency. For surfaces of zero correlation it must always take all values from $+1$ to -1 according to the position of the dividing planes.

We will take one more of these zero correlation tables because it leads up to certain new points. Such a table as that given as Table VIII might well occur in practice; worked out by the product-moment method, the regression lines are linear and there is zero correlation. Table IX shows how extraordinarily the

Diagram III. Frequency surface of zero correlation exhibiting every possible variation of Q with different dichotomic lines.



coefficient of association varies from point to point of division. It is not only zero along the axes, *but zero along a contour line in each quadrant.* Thus, starting from the centre of the surface, we descend in the first and third quadrants to a negative association $-·45$, then crossing the zero contour, we rise to a positive value of $+·50$ and ultimately reach $+·87$! In the second and fourth quadrants the process is reversed; we rise first to $+·45$, then sink to zero and descend first to $-·50$ and

TABLE VIII.

First Member of Pair.

Second Member of Pair.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z		Totals
a′	1	1	2	2	3	4	5	6	6	8	9	11	12	20	12	11	9	8	6	6	5	4	3	2	2	1	1	160
b′	2	2	2	3	3	4	5	8	11	19	31	52	68	130	68	52	31	19	11	8	5	4	3	3	2	2	2	550
c′	2	2	3	3	4	6	8	22	31	69	106	164	252	506	252	164	106	69	31	22	8	6	4	3	3	2	2	1850
d′	2	3	5	8	12	29	41	68	102	133	177	251	389	790	389	251	177	133	102	68	41	29	12	8	5	3	2	3230
e′	3	4	6	15	24	51	79	110	149	191	253	347	532	1056	532	347	253	191	149	110	79	51	24	15	6	4	3	4584
f′	3	7	16	32	60	93	117	161	199	254	326	450	683	1384	683	450	326	254	199	161	117	93	60	32	16	7	3	6186
g′	4	5	25	53	102	128	170	243	257	330	432	588	990	1790	990	588	432	330	257	243	170	128	102	53	25	5	4	8444
h′	4	16	40	99	141	188	232	287	353	441	569	771	1169	2360	1169	771	569	441	353	287	232	188	141	99	40	16	4	10980
i′	4	45	91	152	208	263	327	401	489	610	793	1067	1624	3242	1624	1067	793	610	489	401	327	263	208	152	91	45	4	15390
j′	5	79	161	240	329	411	504	606	751	939	1202	1628	2483	4966	2483	1628	1202	939	751	606	504	411	329	240	161	79	5	23642
k′	20	88	172	263	347	440	531	649	792	1008	1273	1717	2616	5218	2616	1717	1273	1008	792	649	531	440	347	263	172	88	20	25050
l′	5	79	161	240	329	411	504	606	751	939	1202	1628	2483	4966	2483	1628	1202	939	751	606	504	411	329	240	161	79	5	23642
m′	4	45	91	152	208	263	327	401	489	610	793	1067	1624	3242	1624	1067	793	610	489	401	327	263	208	152	91	45	4	15390
n′	4	16	40	99	141	188	232	287	353	441	569	771	1169	2360	1169	771	569	441	353	287	232	188	141	99	40	16	4	10980
o′	4	5	25	53	102	128	170	243	257	330	432	588	990	1790	990	588	432	330	257	243	170	128	102	53	25	5	4	8444
p′	3	7	16	32	60	93	117	161	199	254	326	450	683	1384	683	450	326	254	199	161	117	93	60	32	16	7	3	6186
q′	3	4	6	15	24	51	79	110	149	191	253	347	532	1056	532	347	253	191	149	110	79	51	24	15	6	4	3	4584
r′	2	3	5	8	12	29	41	68	102	133	177	251	389	790	389	251	177	133	102	68	41	29	12	8	5	3	2	3230
s′	2	2	3	3	4	6	8	22	31	69	106	164	252	506	252	164	106	69	31	22	8	6	4	3	3	2	2	1850
t′	2	2	2	3	3	4	5	8	11	19	31	52	68	130	68	52	31	19	11	8	5	4	3	3	2	2	2	550
	1	1	2	2	3	4	5	6	6	8	9	11	12	20	12	11	9	8	6	6	5	4	3	2	2	1	1	160
Totals	80	416	874	1477	2119	2794	3507	4473	5488	6996	9069	12375	19020	37706	19020	12375	9069	6996	5488	4473	3507	2794	2119	1477	874	416	80	175082

TABLE IX.

Values of Mr Yule's Coefficients of Association for Table VIII shown on p. 198.

	a	b	c	d	e	f	g	h	i	j	k	l	m n	o	p	q	r	s	t	u	v	w	x	y	z
a′	+·866	+·634	+·530	+·405	+·343	+·312	+·297	+·283	+·252	+·237	+·212	—	0	—	−·212	−·237	−·252	−·283	−·297	−·312	−·343	−·405	−·530	−·634	−·866
b′	+·811	+·503	+·290	+·131	+·022	−·043	−·084	−·101	−·116	−·097	—	—	0	—	—	+·097	+·116	+·101	+·084	+·043	−·022	−·131	−·290	−·503	−·811
c′	+·636	+·163	−·084	−·256	−·360	−·411	−·442	−·408	−·379	—	—	—	0	—	—	—	+·379	+·408	+·442	+·411	+·360	+·256	+·084	−·163	−·636
d′	+·475	−·046	−·257	−·385	−·448	−·427	−·405	−·344	—	—	—	—	0	—	—	—	—	+·344	+·405	+·427	+·448	+·385	+·257	+·046	−·475
e′	+·388	−·152	−·356	−·427	−·453	−·396	−·342	—	—	—	—	—	0	—	—	—	—	—	+·342	+·396	+·453	+·427	+·356	+·152	−·388
f′	+·300	−·205	−·349	−·379	−·363	−·300	—	—	—	—	—	—	0	—	—	—	—	—	—	+·300	+·363	+·379	+·349	+·205	−·300
g′	+·263	−·299	−·360	−·344	−·288	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	+·288	+·344	+·360	+·299	−·263
h′	+·158	−·298	−·325	−·264	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	+·264	+·325	+·298	−·158
i′	—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—	—	—
j′ k′	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l′	—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—	—	—
m′	−·158	+·298	+·325	+·264	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	−·264	−·325	−·298	+·158
n′	−·263	+·299	+·360	+·344	+·288	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	−·288	−·344	−·360	−·299	+·263
o′	−·300	+·205	+·349	+·379	+·363	+·300	—	—	—	—	—	—	0	—	—	—	—	—	—	−·300	−·363	−·379	−·349	−·205	+·300
p′	−·388	+·152	+·356	+·427	+·453	+·396	+·342	—	—	—	—	—	0	—	—	—	—	—	−·342	−·396	−·453	−·427	−·356	−·152	+·388
q′	−·475	+·046	+·257	+·385	+·448	+·427	+·405	+·344	—	—	—	—	0	—	—	—	—	−·344	−·405	−·427	−·448	−·385	−·257	−·046	+·475
r′	−·636	−·163	+·084	+·256	+·360	+·411	+·442	+·408	+·379	—	—	—	0	—	—	—	−·379	−·408	−·442	−·411	−·360	−·256	−·084	+·163	+·636
s′	−·811	−·503	−·290	−·131	−·022	+·043	+·084	+·101	+·116	+·097	—	—	0	—	—	−·097	−·116	−·101	−·084	−·043	+·022	+·131	+·290	+·503	+·811
t′	−·866	−·634	−·530	−·405	−·343	−·312	−·297	−·283	−·252	−·237	−·212	—	0	—	+·212	+·237	+·252	+·283	+·297	+·312	+·343	+·405	+·530	+·634	+·866

then to $-\cdot 87$. What information can the coefficient of association give us as to the nature of these two *uncorrelated* variates? It is no doubt a measure in some manner of the heteroscedasticity of the arrays. But how and in what way does it measure this phase of want of independence by a value which varies from -1 to $+1$ several times over? If it in some manner measures this heteroscedasticity, it is only by its local values, it measures nothing of the dependence of the variates as a whole. Will Mr Yule tell us how to infer whether, when $Q = +\cdot 87$ or $-\cdot 87$, it is a measure of the relation between the means of the arrays corresponding to given variates, or is merely a measure of the differences in the variabilities of those arrays? Will he also tell us in what manner, by a multiplicity of values, it measures mere heteroscedasticity?

Are we doing Mr Yule an injustice in taking any notice of Q at the extreme limits of our table, e.g. of such values as $+\cdot 81$ or $-\cdot 81$? Well, consider the corresponding fourfold table:

3	707	710
77	174,295	174,372
80	175,002	175,082

$Q = +\cdot 811$

and multiply each entry by the factor

$$\frac{32{,}527{,}843}{175{,}082} = 185\cdot 78633,$$

we find

557	131,351	131,908
14,306	32,381,629	32,395,935
14,863	32,512,980	32,527,843

Compare this with the fourfold table deduced by Mr Yule from the Census data of 1901:

Blindness.

Mental Defect.	Present	Absent	Totals
Present ...	558	132,096	132,654
Absent ...	24,759	32,370,430	32,395,189
Totals	25,317	32,502,526	32,527,843

$Q = +\cdot 693$,

or the following one from the same Census, again formed by Mr Yule:

Deaf-Mutism.

Blindness.	Present	Absent	Totals
Present ...	96	25,221	25,317
Absent ...	15,150	32,487,376	32,502,526
Totals	15,246	32,512,597	32,527,843

$Q = +\cdot 782$

and it will be obvious that we are doing Mr Yule no injustice at all. Our 3 in 175,082 is 557 in 32,527,843 and is far higher than Mr Yule's 96 in 32,527,843! Why his 96 in over 32 millions might almost consist of the persons who had been rendered at the same instant both deaf and blind by accident! Yet on the basis of this result Mr Yule has asserted a "very high association" between deaf-mutism and blindness!

But, since Mr Yule has no hypothesis as to the nature of his frequency, why should not the relation between blindness and deaf-mutism be precisely like that of the nature of the variates exhibited in Table VIII? If this be so, what has Mr Yule's coefficient of association told us? The variates would be actually uncorrelated, but we should anticipate:

(i) That extremely bad sight would be associated with extreme deafness—this is the above $Q = +\cdot 78$.

(ii) Extremely bad sight would be associated with great aural acuity. This has often been asserted of the blind. Here Q would be high and *negative*.

(iii) Extremely good sight would be associated with extremely good hearing, i.e. Q would be high and positive. Persons with exceptionally good capacity of one sense very frequently have it of another sense.

(iv) Extremely bad hearing would be associated with exceptionally good vision, i.e. Q would be high and negative again. This is quite possible, although we have no conclusive evidence on the point. In a small school for deaf-mute girls 90 °/$_\circ$ of the children were found to have normal vision ($\frac{6}{6}$ or $\frac{6}{9}$), none had worse than $\frac{6}{18}$. In a group of children of normal hearing with light hair only 69 °/$_\circ$ had normal vision and 11·5 °/$_\circ$ with vision of $\frac{6}{24}$ or worse. London girls give 85 °/$_\circ$, Glasgow 82 °/$_\circ$ and Edinburgh 80 °/$_\circ$ normal vision—all lower values than in the case of our small sample of deaf mutes.

Results (i)—(iv) would hold if there were no correlation between goodness of sight and hearing—the average sight of a very deaf person being the same as one of normal hearing—provided the variability in sight of the very deaf were less than that of the general population, and the variability in hearing of the very blind were also less than that of the general population*.

Thus given a fourfold table which is based upon continuous variation, if we make no hypothesis with regard to the nature of the frequency, we have in fact no idea at all of what Mr Yule's coefficients of association and colligation really measure. They measure in some form or another deviation from independence, it may be true correlation or it may be heteroscedasticity, and divisions taken at very slight distances apart may give hopelessly divergent values of Q, of which difference of values Mr Yule has given no intelligible interpretation.

* In the case of our Table VIII, the variability of the horizontal character for the whole population $= 1\cdot 6686$, and for the vertical character 1·5979. The variability of the combined two top arrays is 1·6444, and for the extreme vertical column on the right 1·8166. Had the variabilities of the two sets of arrays been the same as those of the general population, the association would have vanished.

If we turn from surfaces of zero correlation to those of finite correlation, we find in the same way that Q takes innumerable values which have no mutual relationship. Heron has already demonstrated this as far as the Gaussian surface of frequency is concerned*. For example, in a Gaussian distribution Mr Yule's association can take every value from ·37 to 1·0 if the correlation be truly ·3 and the divisions be taken *along the diagonal*. These give for a Gaussian distribution the complete range of Q, but it by no means follows that this is true for other types of frequency. Mr Yule, however, makes no hypothesis as to the nature of his frequency surface. To test the kind of meaning Q conveys, suppose the frequency surface to be a rectangular block, length $2a$, breadth $2b$, height h—that is to say, within a given rectangular area the frequency of all combinations of the two variates is equally probable. If the block slopes with its side $2a$ at an angle of θ to the x-variate, we have the correlation

$$r = \frac{(a^2 - b^2)\cos\theta\sin\theta}{\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}\,\sqrt{a^2\sin^2\theta + b^2\cos^2\theta}}.$$

The regression lines are built up of straight lines; for a considerable distance they coincide with the axes of symmetry, but are afterwards bent round horizontally or vertically as the case may be. At the four corners $Q = 0$ and along one pair of parallel sides $Q = +1$, along the other $Q = -1$. Thus two contour lines of zero association pass through the corners pair and pair. In the accompanying diagram (p. 203) the values of Q are given for the special case where $\theta = 45°$, $a = 1{\cdot}5\,b$, and accordingly $r = {\cdot}3846$. Going along the longer axis Q changes from $-1{\cdot}0$ to $-{\cdot}295$ and so to zero, then it rises to $+{\cdot}6$ at the centre and falls to zero again, becoming negative and ultimately concluding with $-1{\cdot}0$ at the boundary. Along the shorter axis Q varies from $+1{\cdot}0$ to $+{\cdot}6$ at the centre and rises to $+1{\cdot}0$ again below it. What can be learnt as to the real association of two variates by a coefficient which behaves in this way? The case would be quite different if Mr Yule had indicated a type of surface for which Q was constant for all divisions and demonstrated that it represented, even with moderate approximation, such distributions as occur in statistical practice. One such surface of "stable association" at any rate is known for the tetrachoric r_t treated as a coefficient of association merely, and that surface is not widely divergent from a considerable number that we actually meet with.

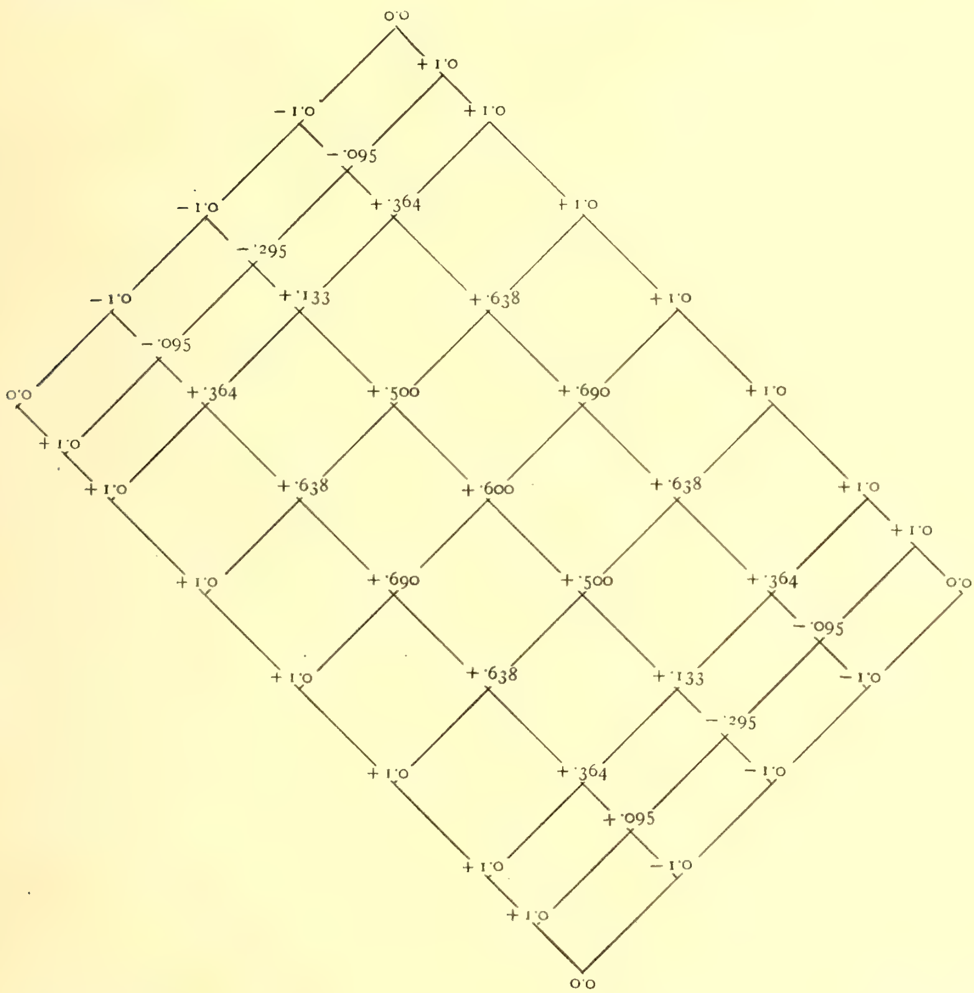
(*b*) *The Fallacy of Mr Yule's Selection in the case of Continuous Variates.*

If Q be not even approximately comparable with itself when taken on the same surface with different dichotomies, how can it be comparable as a measure of any real relationship from one surface to a second? Mr Yule will no doubt reply that a function of Q does measure certain percentages when the table is dressed in an equalised symmetrical form. Our reply is that that form has been obtained by a method of selection which makes very large changes in every other coefficient, including the Boas-Yulean, which has been used to measure

* *Biometrika*, Vol. VIII. p. 109.

relationship; that we dispute entirely the legitimacy of such selection which is of a singularly arbitrary character, both in the extent to which it is applied and the region to which it is circumscribed. If we apply it to continuous frequency surfaces, so that a *certain* Q remains constant, all other Q's are changed, and r taken as a measure of the relationship of the variates as a whole is often

DIAGRAM IV. Q for frequency surface a right six-face. To illustrate how variation of Q depends on form of distribution and how it has no relation whatever to true correlation. Actual correlation ·3846.



immensely changed. A correlation table has a multiplicity of Q's and one product-moment r; a process of selection, which changes all Q's but one, and widely modifies r, has attributed to it by Mr Yule some special merit by which that Q, in preference to any other, is considered for the time being to measure the "association" of the variates!

Any argument, as we have already indicated, which is valid when applied to the columns and rows of a fourfold table ought to be valid when applied to the columns and rows of a multifold table. Such a table should also not be affected by selection. Well, let us take the Table from the Census of the age of husband and wife and let us select so as not to change certain Q's and see what effect this has on the correlation. The following tables give Q and r before and after a series of selections.

Let the divisions be at the same ages for both husband and wife, say under 30 and over 30.

Coefficient of Association $= {\cdot}9745$

Coefficient of Colligation $= {\cdot}7958$

Correlation before Selection $= {\cdot}9136$*

Percentage of those over 30 years selected:

	Actual Correlation		Actual Correlation
100 %	·9136	5 %	·6969
10 %	·8347	4 %	·6479
9 %	·8175	3 %	·5921
8 %	·7961	2 %	·5360
7 %	·7698	1 %	·4952
6 %	·7369	0 %	·4850

In other words the selection which reduces the actual correlation from ·914 to ·485 leaves Mr Yule's coefficients of association and colligation unchanged! No, this is not true; every other coefficient of association and colligation for the table is changed, except the particular two for the arbitrary division at 30 years! What legitimate inference of any kind can be drawn from the *constancy* of this individual pair?

Now let us select husbands and wives unequally but still at age 30 divisions:

Husbands	Wives	Actual Correlation
100 %	100 %	·914
10 % over 30	10 % under 30	·908
1 % over 30	10 % under 30	·850
10 % under 30	1 % over 30	·715
1 % under 30	0·1 % over 30	·285
·000 % under 30	·000 % over 30	− ·009
·000 % over 30	·000 % under 30	− ·038

During all these operations which reduce the actual relationship as measured by correlation from the very high value ·914 to zero and even to negative values, Mr Yule's association and colligation for his selected dichotomies show the constant "very high values" ·975 and ·796. At other dichotomies of course they cover pretty well the whole possible scale.

* Value without Sheppard's corrections, because in dealing with selection it is not clear that those corrections are always appropriate.

Nor has the result anything to do with the division at ages 30. If we divide at 21 years we find:

Percentage selected over 21	Actual Correlation
100 °/。	·914
10 °/。	·939
1 °/。	·941
·1 °/。	·493
·000 °/。	·272

Here the Q and ω of Mr Yule retain throughout the operations the values ·987 and ·853, which mark, we have been told, "very high association"!

Nor is the absurdity in the least confined to ages of husband and wife. Let us take stature in Father and Son* and divide into a fourfold at Fathers 67·5″ and over, and Sons at 68·5″ and over. The actual correlation is ·520, $Q = ·683$ and $\omega = ·395$.

Selection of Father	Selection of Son	Actual Correlation
100 °/。	100 °/。	·520
10 °/。 over 67·5″	10 °/。 over 68·5″	·314
10 °/。 under 67·5″	10 °/。 over 68·5″	·275
·000 °/。 over 67·5″	·000 °/。 over 68·5″	·251
1 °/。 under 67·5″	1 °/。 over 68·5″	·185

The result is exactly the same as before, the real relationship is immensely modified by selection, while the colligation and association remain unchanged for *one* pair of arbitrary divisions and for this one only. What can be learnt from such a statistical method? We venture to believe that from the standpoint of common sense it is wholly without meaning.

What is the precise physical character which is to be attached to this wide difference between "association" and correlation? That correlation is affected by selection we know only too well; it is one of the factors of progressive evolution under natural selection, but what profitable meaning of any kind is to be attached to the statement that one out of an indefinite number of associations has remained unchanged by a special selection? Does not the principle that "association" or "colligation" is unchanged by selection arise from the fact that Mr Yule has merely guessed a denominator to his coefficient, which denominator has no theoretical justification of any kind; and his principle that selection makes no change is a later discovery and has no validity at all, for it is not a "natural" selection and has no generality beyond the fourfold table?

We think we have sufficiently indicated that Mr Yule's coefficients of association and colligation fail entirely for any variates which may be suspected in any way of continuity, and the bulk of the variates to which Mr Yule has applied his methods undoubtedly have such continuity.

* *Biometrika*, Vol. II. p. 415.

Mr Yule would no doubt tell us that he has distinctly stated that he distinguishes between correlation and association and that he knows they may lead to diverse results. We reply that, wherever there is any real continuity, the assumption of a discrete "attribute" disguises its existence and will lead to erroneous conclusions. Further he directly states* that: "The methods applicable to the former kind of observations, which may be termed STATISTICS OF ATTRIBUTES, are also applicable to the latter STATISTICS OF VARIABLES. A record of statures of men for example may be treated by simply counting all measurements as *tall* that exceed a certain limit, neglecting the magnitude of excess or defect, and stating the numbers of *tall* and short (or more strictly not-tall) on the basis of this classification. Similarly, the methods that are specially adapted to the treatment of statistics of variables, making use of each value recorded, are available to a greater extent than might at first sight seem possible for dealing with statistics of attributes. For example, we may treat the presence or absence of the attribute as corresponding to the changes of a variable which can only possess two values, say 0 and 1."

Here Mr Yule directly claims that his methods can be applied to stature, and in the next sentence suggests that it is reasonable to treat the difference between any tall man and any short man as unity because they have been placed under two class-indices "tall" and "short"! He started his statistical work from the standpoint of the pure logician, and he does not perceive that he is applying his reasoning to the class-names of things and not to the things themselves behind these names. Let us take head length and head breadth with a correlation, say, of ·50, and lengths of femur and humerus with a correlation, say, of ·60, then it is perfectly easy by selecting your head length and breadth division and your femur length and humerus length divisions, to make the association between head length and breadth either greater or less than that between femur length and humerus length. What is the value of the coefficient of association as a measure of relationship if this be the case? Every new division gives a different ratio of association for the two sets of attributes. The application of such methods in practice can only tend to the detriment of modern statistical theory.

(8) *On the Application of Mr Yule's Coefficients to Discrete Variates. Mendelism.*

May we not, however, accept Mr Yule's claims for his coefficients of association when the classes differ by a real discrete unit—not a unit arbitrarily introduced by calling a "short" man 0 and a "tall" man 1? Well, the difficulty is to find such cases. However, supposing them to exist, then there is no question that the coefficients of association and colligation are *not* the right methods of approaching the problem, but that the ordinary product-moment correlation coefficient is the

* *Theory of Statistics*, pp. 7—8.

correct method; but this will lead us to results absolutely opposed to those of Mr Yule's association and colligation.

We ourselves, however, doubt the existence of this discrete unit; we have only come across it in *theoretical* Mendelian investigations and doubt whether the "unit character" which is absent or present has any existence in somatic classifications. Of this we should like to give two or three illustrations.

The first illustration we take is from a paper* by Professor E. M. East, entitled "The Mendelian Notation as a Description of Physiological Facts." Professor East is a vigorous Mendelian making the very best defence he can of the Mendelian notation in the present parlous condition of Mendelian theory—which assumes the truth of the unity of the unit. The cross which Professor East cites is one between a "long" corolla and a "short" corolla race of *Nicotiana*. Speaking of other Mendelian investigations into size, Professor East writes of them: "No criticism could be made except that certain of the characters used varied considerably in the mother varieties and therefore were presumably not homozygous for all character factors†. This criticism is apparently answered by a recent investigation of the writer's, as yet unpublished, where two species, *Nicotiana forgetiana* and *Nicotiana alata grandiflora*, were crossed. As seen by the table, the corolla length is very slightly variable in either species‡, nor is it affected to any extent by environment, yet each species was absolutely reproduced by recombination in the F_2 generation."

TABLE X.

Frequency Distributions for Length of Corolla in a Cross between Nicotiana forgetiana (314) *and N. alata grandiflora* (321).

Designation	Class Centres in Millimetres															Totals
	20	*25*	*30*	*35*	*40*	*45*	*50*	*55*	*60*	*65*	*70*	*75*	*80*	*85*	*90*	
314	9	133	28	—	—	—	—	—	—	—	—	—	—	—	—	170
321	—	—	—	—	—	—	—	—	—	1	19	50	56	32	9	167
$(314 \times 321)\ F_1$	—	—	—	3	30	58	20	—	—	—	—	—	—	—	—	111
$(314 \times 321)\ F_2$	—	5	27	79	136	125	132	102	105	64	30	15	6	2	—	828

Now let us call short all "below 60 mm." and long all "above 60 mm." We have all the offspring "below 60 mm." Hence there is "dominance" of short corolla, and we may apply the magic formulae:

$$(DD) \times (RR) = 4\,(DR),$$
$$(DR) \times (DR) = (DD) + 2\,(DR) + (RR),$$

* *The American Naturalist*, Vol. XLVI. p. 639.

† Note the writer's interpretation of the results by the preconceived theory!

‡ The author does not tell us how many plants were grown in each generation. In the parent generation presumably only two. How many were considered in F_1 and F_2? The difference of variability would largely turn on this factor.

with the result that the segregating generation F_2 shows 222 (RR)'s, a very plausible Mendelian quarter.

Prof. East does not do this, although it was a step absolutely compatible with current impressionist classifications by Mendelians. All he says is that "each species was absolutely reproduced by recombination in the F_2 generation" and "I do maintain that the Mendelian notation satisfies the facts of size inheritance as well as it satisfies the facts of qualitative inheritance" (*l.c.* p. 639).

Well, the fourfold Mendelian table of somatic characters would give in the F_2 generation $(RR) = 222$, $(DD) + (DR) = 606$*. Now will anything be discovered by assuming that those two groups differ by a unit? This is the "dichotomy" of Mr Yule's association. We contend that, while *theoretically* (RR) differs from (DR) by having no D as against one D and from (DD) by the latter's having two D's, this theory is idle when pushed into actual Mendelian statistics†. The division at 60 is not a dichotomy of things differing by a unit, except in name, but an arbitrary cut across a continuous distribution, and the application of either the Boas-Yulean ϕ or such coefficients as those of association and colligation is entirely misleading. We are told that "short" is "dominant" over "tall" as a result of experiment, and the importance of dividing at 60 mm. to get our tables is dropped out of sight.

Piebaldism is another unit character of the Mendelians, and doubtless Mr Yule would be content to take his dichotomy between piebald and whole colour. Now here are 2314 mice classified according to the ratio of pigmented area of coat to whole area of coat‡. Where shall we make the division between

Total	*·10, ·15*	*·20, ·25*	*·30, ·35*	*·40, ·45*	*·50, ·55*	*·60, ·65*	*·70, ·75*	*·80, ·85*	*·90, ·95*	*·975*	*1·00*
2314	6	31	60	81	126	111	138	176	79	393	1113

the piebald and the whole colour? The group ·975 largely refers to individuals with a very small white area on belly; but as a matter of fact the 1·00 group has been divided into three sub-groups of individuals who, without having *white* areas on belly, have or have not somewhat lighter pigmented areas there. Where is the true dichotomy, especially when we can show that each grade of piebaldism is hereditary? Mr Yule would no doubt apply his association to piebalds and whole colours as giving a Mendelian unit, but in doing so he will be applying his dichotomy to words, to class-indices and not to the real things represented by them.

* Whether this means anything or not would of course depend on the 105 at 60 never in later generations giving anything below 60!

† Take the division at 35 mm. and we have dominance of long corolla, but the Mendelian quarter now fails.

‡ From the late W. F. R. Weldon's mice data now at press.

Again, an illustration from Mendelian dichotomy may be found in a paper by Hurst entitled "Mendel's 'Law' applied to Orchid Hybrids*." He desired to give a proof that the F_2 generation consists of 50 % of (DR)'s and 50 % of "specifics," (DD)'s and (RR)'s. He recognised that the first cross gave an "intermediate," so he defined his (DD) as all those, which show $\frac{3}{4}(DD)$ character and more, his (RR) as all those that show $\frac{3}{4}(RR)$ character and more, and the "intermediates" or apparently the (DR)'s all those that show character between $\frac{3}{4}(DD)$ and $\frac{3}{4}(RR)$. As a result his "specifics" came out as 2281 and his "intermediates" as 2267 in number, a plausible Mendelian 1 : 1 ratio. Thus the classification into every one of the groups (DD), (DR) and (RR) in the F_2 generation is by trisection of a continuous variate at arbitrary values†.

Pearson has come across an exactly similar instance of the vagueness of the Mendelian unit in breeding dogs. If a short-muzzled dog be crossed with the long-muzzled dog, the hybrid would be described by general impression, and was so considered by him, as short-muzzled. The result was to indicate dominance of the short-muzzle. But when muzzle indices were formed and the dogs' heads measured in a variety of ways, the hybrids were found to be intermediates, and, crossed in again with the short-muzzled stock, they gave a group the mean of which had a position intermediate between the hybrid and that original stock. Each generation had very considerable variation. Dichotomy giving Mendelian ratios was possible, provided an arbitrary division was taken across the continuous distribution. Mr Yule's unit difference, short-muzzle—not short-muzzle, would be a perfectly idle one across what in the F_2 generation is a continuous distribution.

One of the most remarkable features, indeed, of the present situation is the assumption that in some mysterious manner Mr Yule's coefficients of association or colligation can be applied to Mendelian results. Mr Sanger, in his contribution to the discussion, said:

"One additional reason why he welcomed the Paper was that the rise of Mendelian biology had made a great difference. There they were always dealing with things which were discrete, whereas according to all Galtonian laws they always dealt with things which were thought to be continuous. At present there was this difficulty that mathematicians had a prejudice in favour of more elegant mathematics, and the Mendelians had not yet learnt algebra; but that day would come, and then Mr Yule's work would be the work for the Mendelians" (*J. R. S. S.* Vol. LXXV. p. 646).

Mr Yule nowhere repudiated this application of his coefficients, and yet they are the last which can possibly be applied to Mendelian data! We put on one

* *Journal of the Royal Horticultural Society*, Vol. XXVI. Part 4.

† Martin Leake has found continuity in the F_1 generation with an intermediate mean in the case of Indian cottons, *Journal Asiatic Soc. of Bengal*, N. S. Vol. IV. p. 13, 1908. Even more astonishing frequency distributions for the F_2 generation for "talls" and "shorts," number of nodes and lengths of internodes may be obtained from Mr R. H. Lock's "Studies in Plant Breeding in the Tropics," *Annals of Royal Botanic Gardens*, Peradeniya, Vol. II. In these cases it is wholly impossible to speak of a unit difference between the members of either class.

side what Mr Sanger can possibly mean by "all Galtonian laws" dealing with things thought to be continuous—he has clearly never read the treatment of eye-colour in man nor of coat-colour in Bassett hounds by Sir Francis himself, who distinctly treated them as discrete quantities and applies "his laws" to them. But we must indicate the fallacy of applying coefficients of association and colligation to Mendelian characters. The reason for applying them is the assumption made that a Mendelian character is a discrete unit. But if this be so, fourfold, and three by three Mendelian tables should be treated as discrete tables and true product moments formed of them. We believe that one of us was absolutely the first to apply these methods, treating the Mendelian theoretical characters as units*, and his work has been followed up by a whole series of workers in the Biometric Laboratory†. There was thus no question with the Biometric School of how Mendelian *theoretical* problems should be dealt with, and Mr Yule wholly misses the point when he states that Dr Snow's "recent comments in *Biometrika* on the use of the normal coefficient for Mendelian tables in Dr Brownlee's paper" were "a much stronger condemnation of Professor Pearson's than Dr Brownlee's work" (*J. R. S. S.* Vol. LXXV. p. 651). Pearson has never used a normal correlation coefficient on a true fourfold table‡, which he believed to be Mendelian in character. He has only applied such coefficients when he believed the character under consideration to be at bottom continuous, and as far as eye-colour is concerned, the many dissections of eyes he has been able to examine in his recent investigations as to albinism have confirmed rather than weakened that standpoint. But even had he done so, although the use of such a normal coefficient might be criticised on the ground of the labour involved in determining it, it cannot be condemned on any other ground, for it is in all respects as good a coefficient of association as Mr Yule's Q or ω, and possesses the important property that it is subject to selection—in opposition to the wholly fictitious merit which Mr Yule claims for his coefficients, namely that they are uninfluenced by selection.

This point is so well illustrated by Mendelian theoretical tables, that we stay to demonstrate it here. Let us consider the correlation of father and offspring when a population represented symbolically by the fathers

$$l\,(AA) + m\,(Aa) + n\,(aa)$$

is crossed at random with a population of mothers given by

$$l'\,(AA) + m'\,(Aa) + n'\,(aa).$$

Here we can put $N = l + m + n =$ total of fathers, $l' + m' + n' = N' =$ total mothers, and the fundamental Mendelian formulae

$$\begin{aligned}(AA) \times (aa) &= 4\,(Aa),\\ (AA) \times (Aa) &= 2\,(AA) + 2\,(Aa),\\ (Aa) \times (Aa) &= (AA) + 2\,(Aa) + (aa)\end{aligned}$$

* *Phil. Trans.* Vol. 203 A, pp. 53—86, and *R. S. Proc.* Vol. 81 B, p. 225.

† Jacobs, *R. S. Proc.* Vol. 84 B, p 23; Snow, *R. S. Proc.* Vol. 83 B, p. 37, and *Biometrika*, Vol. VIII. p. 420.

‡ The error of Dr Brownlee's work was that he went back on all this and applied continuous methods to *theoretical* Mendelian tables; see *Proc. Roy. Soc. Edin.* Vol. XXX. p. 473.

are assumed to hold. If every possible father be mated with every possible mother so as to insure random mating, we have the contingency table for father and offspring:

Father.

Offspring.	(AA)	(Aa)	(aa)	Totals
(AA)	$2l(2l'+m')$	$m(2l'+m')$	0	$(2l+m)(2l'+m')$
(Aa)	$2l(m'+2n')$	$2m(l'+m'+n')$	$2n(2l'+m')$	$\{2l'(m+2n)+2m'(l+m+n)+2n'(2l+m)$
(aa)	0	$m(m'+2n')$	$2n(m'+2n')$	$(m+2n)(m'+2n')$
Totals	$4l(l'+m'+n')$	$4m(l'+m'+n')$	$4n(l'+m'+n')$	$4(l+m+n)\times(l'+m'+n')$

Now there are four points at which we can make a fourfold classification, α, β, γ and δ, and these will give tables for which the association and colligation in Mr Yule's sense can be calculated. Now we will suppose

$$N = l+m+n \text{ and } N' = l'+m'+n'$$

to remain constant, so that the total population is the same. Then if we divide at α no selection of (AA) or "dominant" fathers will affect the coefficient of association; if we divide at β or δ, Mr Yule's coefficients are unity, or there is perfect relationship between parents and offspring, and if we divide at γ, no selection of recessives will affect the Yulean association. What light can Mr Yule's coefficients possibly throw on Mendelian inheritance, when for two possible divisions they make the parental relationship perfect and for the other two they give substantial values of the relationship, but render it completely independent of selections, which in reality widely influence the relationship of parent and offspring, if we proceed by the theory of discrete units? If we accept—which the present writers do not—the theory of dominance and assert that (AA) and (Aa) are somatically identical, and represent one somatic character, then γ is the only reasonable division to make, and the typical Mendelian table becomes:

Father.

Offspring.	(AA), (Aa)	(aa)	Totals
(AA), (Aa)	$4(l+m)N'-m(m'+2n')$	$2n\{2N'-(m'+2n')\}$	$4NN'-(m+2n)(m'+2n')$
(aa)	$m(m'+2n')$	$2n(m'+2n')$	$(m+2n)(m'+2n')$
Totals	$4(l+m)N'$	$4nN'$	$4NN'$

We have at once from this table Mr Yule's coefficient of association

$$Q = \frac{N'(2l+m)}{N'(2l+m)+m(2l'+m')}.$$

But the correlation r, as found by the product-moment method for discrete unit characters, is

$$r = \frac{\sqrt{n}\sqrt{m'+2n'}\,(2l+m)}{\sqrt{(m+2n)(l+m)\{4NN'-(m+2n)(m'+2n')\}}}.$$

Now examine Mr Yule's coefficient:

(i) There is perfect association, $Q=1$, when we take no heteromorphic fathers $(m=0)$.

The value of r $$= \sqrt{\frac{l(m'+2n')}{2lN'+n(2l'+m')}},$$

and takes, as it should do, all sorts of values according to the nature of the mothers, and the proportions of true dominant and true recessive fathers in the whole paternity. Thus, for all fathers recessive, it is zero, and for all fathers true dominant $(n=0)$ with all mothers purely recessive $(l'=m'=0)$ it is unity. Thus the true correlation under the influence of selection can take every possible value while Mr Yule's coefficient gives perfect association throughout!

(ii) There is perfect association, $Q=1$, when we take only recessive mothers, i.e. $l'=m'=0$.

The value of r $$= \sqrt{\frac{n(2l+m)}{(m+2n)(l+m)}},$$

and depends entirely on the distribution of fathers. When there are no heteromorphic fathers $(m=0)$, it is unity. When there are no recessive fathers $(n=0)$, it is zero. That is to say, while Mr Yule's coefficient shows perfect relationship throughout, the true correlation or real association can run through the whole range from zero to unity.

It is not too much to say that those who suggest that Mr Yule's coefficients of association and colligation will be of service in Mendelian problems cannot have had any acquaintance with the nature of those problems at all. Mendelian *theory* relates to discrete units and every coefficient which is uninfluenced by selection is on that very ground wholly unsuited for use with such units. Selection modifies correlation when we deal with discrete units just as much as when we deal with continuous characters, and any coefficient is valueless which directly starts with the property that it will not be modified by selection.

Mendelian *practice* classifies under unit designations individuals which, as we have just indicated, often show no sharp line of division at all. In such cases to treat the difference of two classes as a unit is juggling with class names, not dealing with the things so classed. Mr Yule has failed in these matters because

he starts from the field of pure logic and not from the observation and record of actualities. Even if the actual Mendelian differences were units, not the differences of continuous variation, then ϕ would be the right coefficient to use, not those of colligation or association*. But even here the results will be often difficult to interpret. In the usual case, however, of Mendelian practice, what we need is not the value of a correlation, but an investigation of whether observation is a reasonable fit to theory, i.e. we must use the ordinary "Goodness of Fit" test†. This point is discussed in Appendix II, as there has recently been some misinterpretation of the matter.

* We have taken the series of those symmetrical fourfold tables for which ϕ has always the Mendelian value 1/3; the values of Q range from ·6 to 1. What interpretation can association give of such Mendelian tables?

† The evil done by Mr Yule's preaching of association to the neglect of more general methods is manifest in a recent paper by G. N. Collins in *The American Naturalist*, Vol. XLVI. p. 572. He gives such numbers as the following for flower colour and long pollen in hybrid sweet peas, taken from Bateson, Saunders and Punnett, *Report III to the Evolution Committee*, p. 9, 1906. The calculated

	Purple		Red		White	
	Long	Round	Long	Round	Long	Round
Observed ...	1528	106	117	381	1199	394
Calculated ...	1448·5	122·7	122·7	401·5	1220·5	407·4

numbers are curious; the authors do not explain adequately how they have obtained them. Assuming them to be correct—but of this we have doubts—the problem proposed by Mr Collins is to determine whether the observed "Purple" and "Red" as distributed into "Long" and "Round" are a random sample from the calculated values, i.e. we compare

Observation ...	1528	106	117	381
Theory	1448·5	122·7	122·7	401·5

Mr Collins remarks "No method has been proposed for making definite comparisons between such series of numbers" (p. 572), and continues "A customary and direct method of comparing the degree of relationship that exists between any two characters is to compute the coefficient of correlation or Yule's 'coefficient of association.'" In the discussion which follows Mr Yule's "coefficient of association" (1900) is used. Considering the work of Pearson and Elderton on "Criteria of Goodness of Fit" (*Phil. Mag.* Vol. L. 1900, pp. 157—175, and *Biometrika*, Vol. I. pp. 155—163), Mr Collins can hardly have gone far in statistics, for how would he have proceeded had he included the "White" in his series? The proper method appears to us the general one, i.e. to determine the probability P of the recorded divergence between observation and theory, calculating

$$\chi^2 = \text{sum}\,\frac{(\text{observation} - \text{theory})^2}{\text{theory}},$$

and deducing P by Elderton's Tables: see our Appendix II. In this case a deviation as large as that observed would only occur once in twenty-one trials or the odds are 20 to 1. But we believe Messrs Bateson and Punnett have done themselves injustice. We do not write this in disparagement of Mr Collins' work; he is undoubtedly right in demanding some test for "goodness of fit" in these luxuriant Mendelian formulae.

(9) *On the Limitation in Value of the Boas-Yulean ϕ.*

Given two total variate frequencies, if we can assert nothing of the nature of the distribution, the maximum value of the uncorrected mean square contingency coefficient depends on the number of cells and cannot for a finite number of cells exceed a certain limit. Mr Yule has spoken of this fact as if it were a serious blot on the method of contingency. We do not agree with him, but it is singular that if he thinks so, he should not have rejected the use of ϕ, the "theoretical value of the correlation." The fact that ϕ had a maximum limit was known to Mr Yule*, yet he never throughout his paper refers to it as detrimental to his own "theoretical value of the correlation." Consider any table:

$\frac{n_1 m_1}{N} + x$	$\frac{n_1 m_2}{N} - x$	n_1
$\frac{n_2 m_1}{N} - x$	$\frac{n_2 m_2}{N} + x$	n_2
m_1	m_2	N

This is the most general form the fourfold can take for given n_1, n_2, m_1, m_2. We then have

$$\phi = xN / \sqrt{n_1 n_2 m_1 m_2}.$$

This will be a maximum of a positive kind when x takes the largest possible value, i.e. when x is equal to the lesser of $n_1 m_2/N$ and $n_2 m_1/N$. It will be a maximum of negative kind when x is equal to the lesser of $\frac{n_1 m_1}{N}$ and $\frac{n_2 m_2}{N}$. Thus ϕ always lies between definite limits which may be most restricted.

Consider the table

$2269 + x$	$97261 - x$	99530
$11 - x$	$459 + x$	470
2280	97720	100000

Here the limits are given by $x = 11$ and $x = -459$ or we have the two tables

2280	97250	99530
0	470	470
2280	97720	100000

and

1810	97720	99530
470	0	470
2280	97720	100000

These give $\phi = \cdot 0106$ and $\phi = -\cdot 4499$.

* *Journal of R. S. Soc.* Vol. LXXV. p. 604.

Mr Yule's association coefficient is for the two cases

$$Q = 1 \quad \text{and} \quad Q = -1.$$

In the first case the tetrachoric r_t is $+1\cdot0$ and in the second tetrachoric $r_t = -1\cdot0$, since $h = 2\cdot0$ and $k = 2\cdot6$*. As another illustration take

5000	0	5000
4772	228	5000
9772	228	10000

This table gives the least positive ϕ for the given marginal frequencies; and with these frequencies ϕ can never rise above the value $+\cdot1527$ which it has for this table. But both Q and tetrachoric $r_t = +1\cdot0$.

The minimum value of ϕ for the same total frequencies is given by the table

4772	228	5000
5000	0	5000
9772	228	10000

This has $\phi = -\cdot1527$, while $Q = -1$ and tetrachoric $r_t = -1$. In other words, ϕ is restricted to lie between $+\cdot1527$ and $-\cdot1527$, while Q and tetrachoric r_t may pass through the whole range $+1\cdot0$ to $-1\cdot0$. Why has Mr Yule not pointed out these facts when recommending ϕ as obviously suitable for all fourfold tables when we get rid of normal variation? Clearly, if he had done so his criticism of *uncorrected* contingency would have been shown to apply with still greater force to his own coefficient. The standard deviations of the variates of the fourfold table $\frac{a \mid b}{c \mid d}$ are *not* given by $\frac{\sqrt{(a+c)(b+d)}}{N}\Delta$ and $\frac{\sqrt{(a+b)(c+d)}}{N}\Delta'$ and the product moment by $\frac{ad-bc}{N^2}\Delta\Delta'$, unless we may concentrate each variate into points at distances Δ and Δ' which Mr Yule takes as units. But it is clear that when we have done this we (i) have fixed the standard deviations of the variates, (ii) can shift our dichotomic lines throughout the whole ranges of Δ and Δ' without influencing the result. If the variates are really not concentrated into points their standard deviations are wholly independent of the dichotomic lines, and every shifting of those lines will change the proportions of each variate falling into the two categories. The independence of the standard deviations of the dichotomic lines is the advantage of the tetrachoric r_t over ϕ; and in practical Mendelian statistics it is in most cases impossible to shift the dichotomic lines without

* Actually $h = 1\cdot999087$ and $k = 2\cdot597180$. As we have already indicated and shall further emphasise in Appendix I, the value of tetrachoric r_t is indeterminable by the usual method in such cases.

modifying the frequencies. This is the ground we have had for applying ϕ to theoretical but not to practical Mendelism*.

A further illustration of this limitation of ϕ for given marginal frequencies of the fourfold is provided in the accompanying Diagram V (p. 217). Here the range of values possible for the Boas-Yulean is given for the special case where one variate has a median division, and the percentage at which the dichotomy of the other variate takes place is given on the horizontal line; for example, for a $10\,^\circ/_\circ$ dichotomy ϕ must lie between $\pm\,\cdot3333$. We do not ourselves lay stress on this limitation of the range of values in the Boas-Yulean, but if it be a defect of the coefficient of mean square contingency that for a fourfold table its value cannot exceed ·707, it is also a defect of the coefficient recommended by Mr Yule that it also has a limited range for given marginal frequencies, a limitation not shared by the tetrachoric coefficient or even Mr Yule's coefficient of association.

(10) *The Coefficient of Contingency.*

We do not propose to take up at great length a defence of this coefficient because one of us has had for some years a memoir on the subject in hand which will soon see the light of day. But Mr Yule's criticisms arise from two sources, (i) from his disregard of corrections which practice has taught us were needful and which have been known for some time, (ii) from his obvious want of that confidence in the method which arises from long experience of its applicability.

The corrections needed are (*a*) those due to number of cells, and (*b*) the correction for class-index. If κ = number of rows, λ = number of columns, then on the average of many random samples the correction for number of cells is

* We have the following results for the small-pox data:

	Boas-Yulean ϕ	Possible range of ϕ for given frequencies
Sheffield	·531	+ ·9181 to − ·1221
Leicester	·249	+ ·2806 to − ·2228
Homerton-Fulham	·423	+ ·8101 to − ·2301

How would Mr Yule compare these values of ϕ with each other or with those of r from continuous frequencies, which can range from −1 to +1, or again with a Boas-Yulean ϕ from such tables as

499,200	800	500,000
498,306	1694	500,000
997,506	2494	1,000,000

$\phi = +\cdot02$,

Possible range $+\cdot05$ to $-\cdot05$,

499,988	12	500,000
499,987	13	500,000
999,975	25	1,000,000

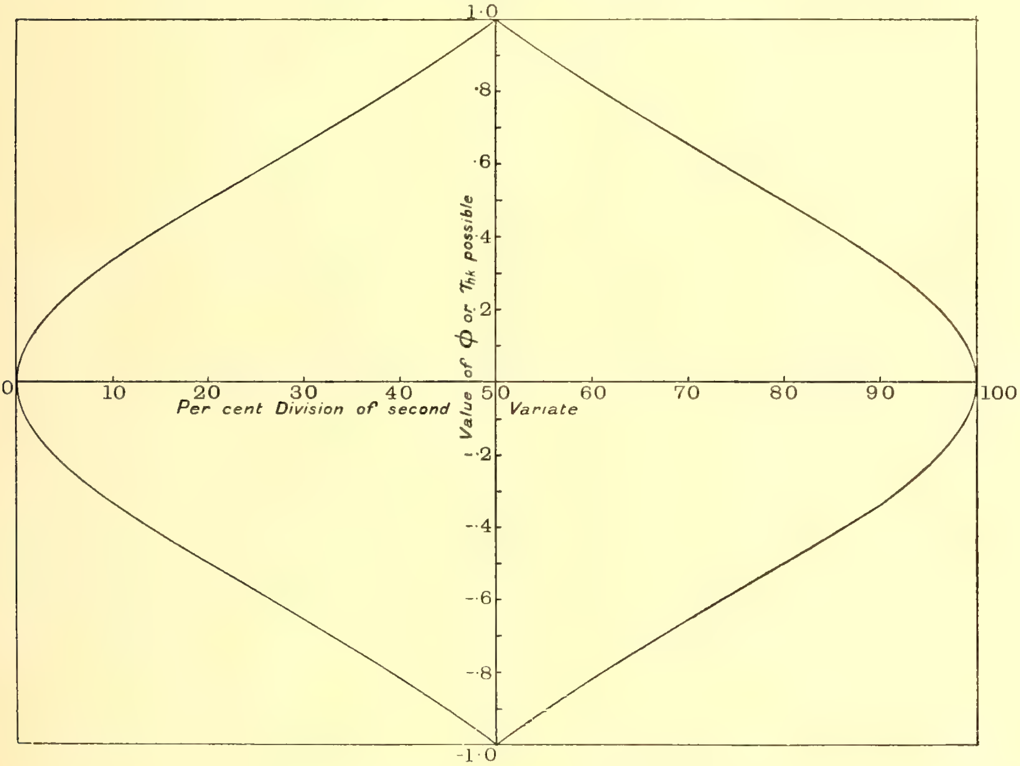
$\phi = +\cdot0003$,

Possible range $+\cdot005$ to $-\cdot005$?

The bulk of the mental defect and blindness data considered by Mr Yule has for ϕ total possible ranges varying from ·4 to ·6 on the positive side and ·006 to ·004 on the negative side. How can the resulting coefficients be intercomparable?

$(\kappa-1)(\lambda-1)/N$ to be subtracted from ϕ^2. This is the chief but is not the only correction for number of cells. It is, however, the one of most importance for our present purpose. It must only be applied when our material can be looked upon as a random sample. It should not be used of course when our material is an actual theoretical frequency surface, and not a random sample from such a surface. The second correction is for the use of class-indices in grouping. The theory of this correction is discussed in an earlier paper in this number of *Biometrika* (Vol. IX. p. 116), where it has been detached from the memoir in preparation on contingency in order to indicate certain fallacies in Mr Yule's statistical theories.

DIAGRAM V. Maximum and minimum values of ϕ (or r_{hk}) for a 50 % division of one category and various percentage divisions of the other.



The area *inside* the curved figure contains all the possible values of ϕ.

Each variate must be corrected independently for the use of broad categories, by calculating the correlation of the variate with its class-index. In order to test the efficiency of the coefficient of contingency for a variety of groupings an arbitrary series of groups must first be selected to work upon. We choose the groupings of the eye-colour data published by Pearson and Lee for father and son as being perfectly arbitrary groupings fixed before any controversy arose on this subject

and therefore clearly not selected to obtain favourable or unfavourable results. These frequencies are

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	Totals
Father...	36	322	264	180	5	64	101	28	1000
Son ...	34	301	284	137	5	100	98	41	1000

The reader should remember that the eye-colour groups are as follows:

1 = light blue, 5 = light brown,
2 = blue, dark blue, 6 = brown,
3 = blue-green, hazel, 7 = dark brown,
4 = dark grey, hazel, 8 = very dark brown, black.

Group 5 has been usually clubbed with group 6 because in 5 only five fathers and five sons occur.

The following table gives the class-index correlations for various groupings:

TABLE XI.

Table of Class-Index Correlations.*

Order of Table	Divisions	Class Correction Fathers	Class Correction Sons	Corrective Dividing Factor
1st 7×7	1 : 2+3 : 4 : 5 : 6 : 7 : 8	·9010	·9011	·8118,9110
2nd 7×7	1 : 2 : 3 : 4 : 5+6 : 7 : 8	·9624	·9645	·9282,3480
1st 6×6	1 : 2+3 : 4 : 5+6 : 7 : 8	·9009	·9010	·8117,1090
2nd 6×6	1 : 2 : 3 : 4 : 5+6 : 7+8	·9542	·9553	·9115,4726
1st 5×5	1 : 2+3+4 : 5+6 : 7 : 8	·8070	·8352	·6740,0640
2nd 5×5	1+2 : 3 : 4 : 5+6 : 7+8	·9259	·9296	·8607,1664
1st 4×4	1 : 2+3+4 : 5+6 : 7+8	·7971	·8246	·6572,8866
2nd 4×4	1+2 : 3+4 : 5+6 : 7+8	·9054	·9144	·8278,9776
3rd 4×4	1+2 : 3 : 4 : 5+6+7+8	·9156	·9130	·8359,4280
1st 3×3	1+2+3 : 4 : 5+6+7+8	·8253	·8191	·6760,0323
2nd 3×3	1+2 : 3+4 : 5+6+7+8	·8949	·8975	·8031,7275
3rd 3×3	1 : 2+3+4+5+6+7 : 8	·5669	·5968	·3383,2592

The reader should note that while the corrective dividing factor on the whole gets smaller and smaller as the classes get fewer and fewer, yet it is not possible to assert *a priori* that more classes will have a higher factor than fewer classes, e.g. our 1st 7 × 7 table has a lower corrective factor than our 3rd 4 × 4 table, and our 2nd 3 × 3 table a higher factor than our 1st 5 × 5 table. This point is important in reference to a criticism we shall make later of Mr Yule's statistical methods.

* The justification for obtaining the class-index correlation by determining the mean by the Gaussian hypothesis is given later in this paper; see also *Biometrika*, Vol. IX. pp. 127, 139, etc.

Taking this system of classification we divided up a Gaussian surface of ·5 correlation into the same groups, and also a Gaussian surface of ·3 correlation. We publish these surfaces below*.

TABLE XII.

Gaussian Surface for $r = \cdot 5$ in Eye-Colour Groupings.

	1	*2*	*3*	*4*	*5+6*	*7*	*8*	Totals
1	7·38	19·85	4·94	1·38	0·26	0·18	0·01	34
2	20·58	145·47	78·94	35·98	9·72	9·27	1·04	301
3	6·01	93·63	85·41	54·34	18·59	22·33	3·69	284
4	1·26	31·81	39·49	31·03	12·29	17·36	3·76	137
5+6	0·53	18·11	27·79	25·14	11·09	17·62	4·72	105
7	0·22	11·02	21·59	23·66	11·86	21·89	7·76	98
8	0·02	2·11	5·84	8·47	5·19	12·35	7·02	41
Totals	36	322	264	180	69	101	28	1000

TABLE XIII.

Gaussian Surface for $r = \cdot 3$ in Eye-Colour Groupings.

	1	*2*	*3*	*4*	*5+6*	*7*	*8*	Totals
1	4·04	17·16	7·55	3·30	0·91	0·92	0·12	34
2	17·41	123·59	79·76	44·64	14·61	17·67	3·32	301
3	8·86	93·00	78·31	52·04	19·20	26·40	6·19	284
4	2·83	37·73	37·24	27·51	10·95	16·31	4·43	137
5+6	1·62	25·21	27·75	22·09	9·26	14·64	4·43	105
7	1·02	19·50	24·47	21·39	9·58	16·36	5·68	98
8	0·22	5·81	8·92	9·03	4·49	8·70	3·83	41
Totals	36	322	264	180	69	101	28	1000

In applying the method of contingency to these two tables, no correction for mean of ϕ^2 in the random sample should be made; they are actual surfaces and not random samples from these surfaces. Further in order to measure what effect dealing only with round numbers in the cells would make we replaced the *first* table by the following Table XIV in which the decimals were cut off and a slight adjustment made to preserve the total variate frequencies. This table is published so that the reader can judge on what we worked.

This working to units cannot be expected to give quite as good a result as working to two decimal places, but it is more consonant with an actual table.

Next we took Pearson's Family Data cards and arranged 1000 cases of Father and Son, first in order of magnitude of Father's stature, next in order of Son's

* Mr Yule states that he has divided up the ·3 Gaussian surface in a somewhat similar manner, but he does not publish his table, and it is therefore impossible to test his results. We should like here to enter a protest against this procedure, which recurs in Mr Yule's memoir, and throws an immense amount of unnecessary arithmetic on any one traversing Mr Yule's arguments.

TABLE XIV.

Gaussian Surface for $r = \cdot 5$ adjusted to give whole units in cells.

	1	*2*	*3*	*4*	*5+6*	*7*	*8*	Totals
1	7	20	5	2	—	—	—	34
2	21	145	79	36	10	9	1	301
3	6	94	85	54	19	22	4	284
4	2	32	39	31	12	17	4	137
5+6	—	18	28	25	11	18	5	105
7	—	11	22	24	12	22	7	98
8	—	2	6	8	5	13	7	41
Totals	36	322	264	180	69	101	28	1000

magnitude of stature and divided up the frequencies of Father's and Son's stature exactly in the eye-colour groups. The following contingency table resulted.

TABLE XV.

Stature of Father and Son in Eye-Colour Groups.

Stature of Father.

Stature of Son.		*1*	*2*	*3*	*4*	*5+6*	*7*	*8*	Totals
	1	4	22	7	—	1	—	—	34
	2	23	154	84	26	8	6	—	301
	3	8	87	75	66	22	24	2	284
	4	1	29	36	37	14	14	6	137
	5+6	—	18	27	26	11	18	5	105
	7	—	9	26	19	7	29	8	98
	8	—	3	9	6	6	10	7	41
	Totals	36	322	264	180	69	101	28	1000

We have thus one table which for practical purposes is absolutely Gaussian, one Gaussian table modified to give units in the cells, and one table typical of what occurs in, perhaps, 9 out of 10 cases in every-day statistics.

The next table gives the results obtained by the method of mean square contingency with the appropriate corrections. It will be seen that 3×3 tables give as good results as 7×7 tables and the method is thus justified for Gaussian material and for the bulk of such tables as occur in statistical practice.

It will be seen not only how closely the mean of the contingency values agrees with the product-moment value of the correlation, but how little the individual values differ from the mean. Worse cases may possibly be found by those who go to seek them by extreme divisions—we have taken them as they came, and feel convinced by a wide experience that contingency gives in practice remarkably satisfactory results.

We now take a step further and ask, if we depart from the ordinary run of cases and pick out skew distributions, can we place equal reliance on the contingency

Order of Table	Classes Grouped	Stature Father and Son	Gaussian Surface $r=0{\cdot}5$	Gaussian Surface $r=0{\cdot}3$
2nd 7×7	1 : 2 : 3 : 4 : 5+6 : 7 : 8	·49	·49	·30
2nd 6×6	1 : 2 : 3 : 4 : 5+6 : 7+8	·49	·48	·30
2nd 5×5	1+2 : 3 : 4 : 5+6 : 7+8	·52	·48	·31
2nd 4×4	1+2 : 3+4 : 5+6 : 7+8	·52	·49	·30
3rd 4×4	1+2 : 3 : 4 : 5+6+7+8	·51	·48	·30
2nd 3×3	1+2 : 3+4 : 5+6+7+8	·51	·48	·31
1st 3×3	1+2+3 : 4 : 5+6+7+8	·52	·51	·30
Mean by Contingency	—	·5080*	·4884†	·3017 ‡
Product-Moment Value	—	·5189 §	·5000	·3000
7×7 Product Moment	Classes concentrated at Gaussian means and corrections used¶	·5231	·5023	·3005

method to give results of *practical* value? By practical value we mean results within ·05 of the true value of the correlation, for very small weight is given in practical statistics to deviations of less than this order. We cannot do better in answering this problem than by taking the very surfaces (some of which were originally selected by Pearson to illustrate extreme non-Gaussian material) which Mr Yule has gone out of his way to collect, for they are very far from random samples of average statistical experience.

These cases are (i) a hypothetical Mendelian surface constructed by Pearson and noted by him as skew at the time, (ii) the barometric table for Laudale and Southampton, (iii) the ages of husband and wife, (iv) the length of ivy leaves in various stages of growth—all cases selected as tests by Mr Yule.

The following Tables give the data as we have used them. It will be seen from this material (i) how wide is the divergence from Gaussian type and (ii) what a large range of diverse classifications have been used.

There is here extreme deviation from the Gaussian type, the arrays have every variety of skewness from the J-shaped curve of the zero couplets to the normal symmetry of the four couplets arrays. The actual correlation as found by the product-moment method is $\frac{1}{3}$.

Now in this case there is (i) no corrective factor for random sampling as the Table is a theoretical table and not a random sample from such a table, (ii) there is no correction for class-indices because the class-indices are the actual values.

* The 1st 4×4 and 1st 5×5 were also worked out and gave respectively ·5151 and ·5174.

† The slightly greater divergence from the true value here was we believe due to the adjustment of the table to unit frequencies.

‡ The 1st 6×6 and the 1st 5×5 tables were also worked out and gave ·3049 and ·3089 respectively.

§ The original table of Father and Son with 1078 entries gave $r=\cdot5140$; see *Biometrika*, Vol. II. p. 378.

¶ Class-index corrections made: see *Biometrika*, Vol. IX. p. 128.

We have $\phi^2 = \cdot1251373$ and $C_2 = \cdot3335$ against actual $r = \cdot3333$. A better illustration of the value of C_2 properly used could hardly be imagined.

TABLE XVI.

Mendelian Inheritance of Recessive Couplets on Pearson's Theory.*
Case of Four Couplets.

Number of Recessive Couplets in Parent.

Number of Recessive Couplets in Off-spring.	0	1	2	3	4	Totals
0	625	500	150	20	1	1296
1	500	800	360	64	4	1728
2	150	360	276	72	6	864
3	20	64	72	32	4	192
4	1	4	6	4	1	16
Totals	1296	1728	864	192	16	4096

We now take:

TABLE XVII.

4 × 4 *Table for Homotyposis in Ivy Leaves*†.

First Leaf.

Second Leaf.	(1) Under *6·95*	(2) *6·95—10·95*	(3) *10·95—14·95*	(4) Over *14·95*	Totals
(1) Under *6·95*	3358	3133	452	41	6984
(2) *6·95—10·95*	3133	13566	7497	1124	25320
(3) *10·95—14·95*	452	7497	10474	2889	21312
(4) Over *14·95*	41	1124	2889	2330	6384
Totals	6984	25320	21312	6384	60000

In the original paper Pearson gives the correlation as ·562 ± ·009. Reworking the original table with use of Sheppard's correction it is found to give ·567. The Table is extremely skew, and owing to its method of construction—i.e. taking pairs of leaves out of 25 gathered from each of 100 plants—it is markedly lumpy. The massing into a 4 × 4 table has removed this lumpiness and the result probably represents the organic relation between leaves of ivy from the same spray better than the original table. We find:

$\phi^2 = \cdot353{,}307$ when corrected for random sampling, $C_2 = \cdot51095$, and the class-index correlation given by $r^2{}_{C_x x} = \cdot870{,}404$, which leads to ·587 for the corrected contingency. No investigation has been made of the linearity of the regression, but it seems to us that the value of the correlation in its general sense may be

* *Phil. Trans.* Vol. 203 A, p. 60. Here is another illustration of Mr Yule's peculiar methods. This table takes some calculating, but Mr Yule does not give the table, so that his results might be verified. Further any reader of his memoir would suppose that Pearson had applied on this occasion or on other occasions tetrachoric r_t to tables of this character, whereas the exact contrary is the fact, Pearson having been the first to apply product-moment formulae to *theoretical* Mendelian data, including this case.

† *Phil. Trans.* Vol. 197 A, p. 351.

even better given by the ·587 of the contingency method than by the ·567 of the product-moment method. Anyhow the difference is of no practical significance; and we again see that contingency applies effectively to skew material.

We next take into consideration a 7 × 7 table, the barometer data from Laudale and Southampton.

TABLE XVIII.

7 × 7 *Table for Barometer Heights at Laudale and Southampton.*

Southampton.

Laudale.	(1) Over *30·55*	(2) *30·55—30·15*	(3) *30·15—29·95*	(4) *29·95—29·75*	(5+6) *29·75—29·65*	(7) *29·65—29·35*	(8) Below *29·35*	Totals
(1) Over *30·45*	50	63·75	2·25	—	—	—	—	116
(2) *30·45—30·05*	42	503·25	248·5	64·5	6·25	—	—	864·5
(3) *30·05—29·75*	—	193·5	340	221·25	39·5	41·75	2·5	838·5
(4) *29·75—29·55*	—	35	120·25	169	45	72·75	15	457
(5+6) *29·55—29·35*	—	4·5	49·5	117·75	63·25	77·25	14·75	327
(7) *29·35—28·95*	—	—	17·5	54	46	102·25	46·75	266·5
(8) Below *28·95*	—	—	—	1	1	20	30·5	52·5
Totals	92	800	778	627·5	201	314	109·5	2922

This is a singularly unfavourable table for contingency methods for it is a well-known rule in practical working to avoid cells whose actual frequencies or those of independent probabilities are zero or a few units. We should therefore anticipate getting the best results in such cases from few divisions in which cells with zero or small entries rarely occur. We have numbered the divisions to correspond with the eye-colour data nomenclature.

Order of Table	Divisions	Mean Square Contingency	Southampton r_{xC_x}	Laudale r_{yC_y}	r_{xy}
7×7	1 : 2 : 3 : 4 : 5+6 : 7 : 8	·69791	·9667	·9659	·75
6×6	1 : 2 : 3 : 4 : 5+6 : 7+8	·67755	·9570	·9596	·74
5×5	1+2 : 3 : 4 : 5+6 : 7+8	·63210	·9345	·9324	·73
1st 4×4	1+2 : 3+4 : 5+6 : 7+8	·60952	·9089	·9136	·73
2nd 4×4	1+2 : 3 : 4 : 5+6+7+8	·61368	·9238	·9164	·72
1st 3×3	1+2 : 3+4 : 5+6+7+8	·58899	·8979	·8973	·73
2nd 3×3	1+2+3 : 4 : 5+6+7+8	·55020	·8450	·8215	·79
3rd 3×3	Extra*	·57530	·8823	·8736	·75
4th 3×3	Extra*	·45500	·7972	·7540	·76
Data classified to ·1″ give Product Moment r = ·780†				Mean	·744

* Two cases of this for 3 × 3-fold tables had already been worked out by Pearson in his paper in the earlier part of this number and are cited here; see *Biometrika*, Vol. IX. pp. 136—7.

† Given as ·757 in the original paper, *Phil. Trans.* Vol. 190 A, p. 455, which antedated the publication of the correct 'Sheppard's' corrections.

Here again for practical purposes the mean square contingency gives quite good results. It is very unlikely that in practical statistical work stress would be laid on a difference of ·05 occurring in a correlation of this magnitude.

We now turn to the table of ages of Husband and Wife taken from the Census of 1901. We have first arranged an 8×8-fold table in unequal frequency groups, as follows:

TABLE XIX.

Age of Wife.

Age of Husband.	(1) *15—20*	(2) *21—29*	(3) *30—39*	(4) *40—44*	(5) *45—49*	(6) *50—54*	(7) *55—64*	(8) *65—*	Totals
(1) *15—24*	44112	194620	4962	162	40	16	8	1	243921
(2) *25—34*	15763	881320	590180	14046	2669	645	186	17	1504826
(3) *35—44*	873	98200	910238	388586	78648	13986	3362	240	1494133
(4) *45—49*	90	6584	85583	178474	251519	58699	12223	576	593748
(5) *50—54*	43	2573	27000	56976	146990	194509	53322	2027	483440
(6) *55—59*	29	1166	10275	17810	46192	110115	176506	7847	369940
(7) *60—69*	16	788	6570	10233	21314	50238	261167	102356	452682
(8) *70—*	7	184	1257	1780	3290	6759	36251	125302	174830
Totals	60933	1185435	1636065	668067	550662	434967	543025	238366	5317520

The following cases were investigated by contingency:

Order of Table	Division	Contingency uncorrected for classes	Wife $r_{C_x x}$	Husband $r_{C_y y}$	r_{xy}
6×6	1+2 : 3 : 4 : 5 : 6 : 7+8	·7821	·9470	·9346	·88
5×5	1+2 : 3 : 4 : 5+6 : 7+8	·7677	·9447	·9327	·87
4×4	1+2 : 3 : 4+5 : 6+7+8	·7405	·9312	·9223	·86
3×3	1+2 : 3+4+5 : 6+7+8	·6995	·8959	·8954	·87

The true value of the coefficient of correlation by product-moment method = ·925. These values in an extreme case of skewness, rendered more complex by heterogeneity of the material*, so far from discrediting the method of contingency are definitely in its favour. On material given by a 3×3-fold or 4×4-fold table no reasoning is likely to be based which would be modified by a change of ·03 to ·05 in the correlation. We think that for even these exceptional cases—which rarely occur without warning in practical statistics—the corrected contingency will not lead any one astray†.

* Skewness of wife's ages = ·71; skewness of husband's ages = ·76. The heterogeneity depends on the fact that the table represents the ages of *all* married couples without regard to second or later marriages.

† Here are a few comparisons of results reached by the method of pseudo-ranks with those deduced by contingency:

What has Mr Yule to place against this method for tables 3 × 3-fold up to 6 × 6-fold in classification? He writes: "From several trials—more than are here given—I have come to the tentative conclusion that the best guide to the correlation that would be found for given data, if the grouping were other than that which in fact it is, is the correlation for the existing grouping, provided that you are given at least some five or six arrays" (*loc. cit.* p. 618). He does not venture to inform us what he would do for a table of 3 × 3, or 4 × 4 cells! As a matter of fact such tables can give even by Mr Yule's method of pseudo-ranks better results than the 5 × 5 or 6 × 6 groupings. Mr Yule is, however, not content with his statement that his method is *tentative*; before he has done with it* he has assumed that by increasing his classes he will approach a limit which is the true product-moment correlation. As a matter of fact there is no such approach at all; the Yulean method of pseudo-ranks may give a better result for a lower than a higher number of cells, and if it did go far enough to reach a limit, it would

Order of Table. Number of Classes	Barometer Results		Age of Husband and Wife		0·5 Gaussian Correlation arranged in Father and Son Eye-Colour Groups	
	Contingency	Pseudo-Ranks	Contingency	Pseudo-Ranks	Contingency	Pseudo-Ranks
7 × 7	·75	·73	—	—	·49	·46
6 × 6	·74	·73	·88	·89	·48	·45
5 × 5	·73	·72	·87	·88	·48	·44
1st 4 × 4	·73	·67	·86	·86	·49	·42
2nd 4 × 4	·72	·71	—	—	·48	·43
1st 3 × 3	·73	·66	·91	·83	·48	·41
2nd 3 × 3	·79	·65	·84	·81	·51	·40
3rd 3 × 3	·75	·67	—	—	—	—
4th 3 × 3	·76	·52†	·87	·79	—	—
Mean	·744	·674	·890	·845	·488	·431
True r	·780	·780	·925	·925	·500	·500

It will be seen that the method of contingency, especially with few classes, is markedly better than that of pseudo-ranks. We have purposely introduced the Age of Husband and Wife, because the divisions there have ranges not very diverse in magnitude. In such cases the method of pseudo-ranks becomes almost exactly that of the true correlation of variates, if the proper Sheppard's correction be made. In the case of Ages of Husband and Wife, *if this correction be included*, we are practically finding the true correlation by the method of pseudo-ranks. It is remarkable that Mr Yule has not drawn attention to this, because it at once indicates how fallacious the method is, if the subranges be unequal.

* On the basis of what he starts by calling a tentative method, he then proceeds to assert that all the biometric pigmentation work is "wholly untrustworthy" (*loc. cit.* p. 622)—a characteristic illustration of how Mr Yule's mind rapidly grows obsessed by a theory which he has not properly investigated.

† This case is of remarkable interest as indicating the futility of Mr Yule's method of pseudo-ranks. The table although 3 × 3-fold has sensibly *equal* ranges. Therefore the correction to pass from ranks to variates is closely Sheppard's. The Yulean pseudo-ranks coefficient thus corrected is raised from ·522 to ·775, close to the true correlation!

give the correlation of ranks and not that of true variates*. The fallacy of Mr Yule's arguments and the extreme inferiority of his method to that of contingency will be manifest in the following comparisons. We, however, draw the attention of the reader to this: that if the method of pseudo-ranks did approach a limit it would be the correlation of ranks uncorrected for huge brackets, i.e. we should still have to correct for passing from ranks to variates and for class-indices.

Comparison of Method of Contingency with Method of Pseudo-Ranks.

Coefficients deduced from Table XV, Stature of Father and Son. Product-Moment Correlation = ·52.

Order of Table	Nature of Divisions	Pearson's Contingency	Yule's Pseudo-Ranks
7×7	1 : 2 : 3 : 4 : 5+6 : 7 : 8	·49	·48
6×6	1 : 2 : 3 : 4 : 5+6 : 7+8	·49	·48
1st 5×5	1 : 2+3+4 : 5+6 : 7 : 8	·52	·37
2nd 5×5	1+2 : 3 : 4 : 5+6 : 7+8	·52	·47
1st 4×4	1+2 : 3 : 4 : 5+6+7+8	·51	·46
2nd 4×4	1 : 2+3+4 : 5+6 : 7+8	·52	·36
3rd 4×4	1+2 : 3 : 4 : 5+6+7+8	·51	·46
1st 3×3	1+2+3 : 4 : 5+6+7+8	·52	·37
2nd 3×3	1+2 : 3+4 : 5+6+7+8	·51	·44

The inferiority of the method of pseudo-ranks will be obvious. The contingency gives as good results for a 3 × 3 table as for a 5 × 5 table†; but for two different tables of the same order the method of pseudo-ranks will give results differing by as much as ·10, ten times the difference of the contingency method.

Here is another Table‡, to which Mr Yule has applied his method of reaching a limit to the actual correlation, namely that for eye-colour for pairs of brothers.

Mr Yule having so to speak ingeniously "dressed the window" to show a falling correlation of pseudo-ranks with his increase of classes, then asserts that this pseudo-rank correlation approaches a limit below 0·28, and extending his fallacious reasoning holds that this limit of ranks is the limit to the true correlation of

* The fog in Mr Yule's mind on this subject is well illustrated by his table on p. 619. He takes a table for a Gaussian distribution of correlation 0·3 and says that with an infinite number of classes the Yulean coefficient would become 0·3. There is not a trace of any knowledge on his part that the limit of a process by which *unit-range* is given to each individual is not the same as that of a process by which *unit-area* of the frequency curve is given to each individual. Here and elsewhere he makes no distinction between the correlation of variates and the correlation of ranks. In the actual case his limit would have been ·2876 and not ·3000, but far greater differences will arise if the material be skew. We return to this point later.

† Like all statistical methods, that of contingency must be used with due regard to the data to which it is applied and to the manner in which it is applied. Compound or heterogeneous material may give a contingency coefficient differing considerably from true correlation, and groupings of great inequality in the cells may render idle the corrective factor, e.g. if the great bulk of the material be placed in one or two cells.

‡ *Phil. Trans.* Vol. 195 A, p. 140.

TABLE XX.

First Brother.

Second Brother.	1	2	3	4	5+6	7	8	Totals
1	16	38	19	10	3	6	6	98
2	38	404	205	53	65	41	27	833
3	19	205	418	97	56	78	28	901
4	10	53	97	168	47	50	18	443
5+6	3	65	56	47	70	42	14	297
7	6	41	78	50	42	72	8	297
8	6	27	28	18	14	8	30	131
Totals	98	833	901	443	297	297	131	3000

variates. Now it is clear that there is no such general rule about the correlation of pseudo-ranks always moving in one direction. It is possible within certain limits to vary that correlation in an almost endless manner according to where we take our divisions. It is far more influenced by the size and position of our "brackets" than by whether we work with a 3×3-fold or a 7×7-fold classification. We can choose in this case a 3×3-fold table to give this spurious coefficient of

Resemblance of Eye-Colour in Brothers.

Order of Table	Nature of Divisions	Pearson's Contingency	Yule's Pseudo-Ranks
7×7	1 : 2 : 3 : 4 : 5+6 : 7 : 8	·51	·29
1st 6×6	1+2 : 3 : 4 : 5+6 : 7 : 8	·52	·29
2nd 6×6	1 : 2 : 3 : 4 : 5+6 : 7+8	·49	·29
5×5	1+2 : 3 : 4 : 5+6 : 7+8	·50	·30
1st 4×4	1+2 : 3 : 4 : 5+6+7+8	·51	·33
2nd 4×4	1+2+3 : 4 : 5+6 : 7+8	·53	·27
3rd 4×4	1+2 : 3+4 : 5+6 : 7+8	·44	·27
1st 3×3	1+2+3 : 4 : 5+6+7+8	·54	·29
2nd 3×3	1+2 : 3+4 : 5+6+7+8	·44	·30
3rd 3×3	1+2 : 3 : 4+5+6+7+8	·50	·36
Mean	... ...	·498	> ·28*

Mr Yule any value from ·19 to almost ·40. This flows from the fact that the class-index correction may take a wide range of values according to the arrangement of the classes, and Mr Yule makes no allowance whatever for this fact†.

* "Can we have any hesitation in similarly estimating the correlation for the eye-colour table, if we were in a position to adopt a finer and more uniform grouping (without assuming that we will compel that grouping to give us a normal distribution) as something slightly *less* than 0·28?" Yule, *loc. cit.* p. 619.

† It is not possible to correct the Yulean pseudo-ranks correlation (i) for passing from ranks to variates, because Mr Yule not only rejects any appeal to the Gaussian, for which we know the proper correction, but because his assumption of unit ranges precludes the use of that curve, nor (ii) for class-index correlation, because the same assumption hinders any rational method of finding the class-index correlations.

We have seen that the corrected contingency gives a good practical approach to the actual correlation, even if the material be skew as in the Husband and Wife or in the Barometric Data; are we to argue from the entirely fallacious reasoning of Mr Yule that the method is in this case on the average 44 °/₀ in error although in the skew Mendelian table it is only ·06 °/₀ in error and in the Husband and Wife data only 3·8 °/₀ in error? We shall need some far better reasons for believing that the value usually assumed for the correlation of eye-colour in brothers, i.e. circa ·50, is in error, than such as Mr Yule seems able to adduce.

Another piece of remarkable special pleading on the same lines is that provided by Mr Yule in the case of Pearson's tables for parental heredity in coat-colour in horses*. Those tables are remarkable in their nature, because although 16 classes are formed, there are practically no entries except in the three main groups Brown, Bay and Chestnut. Here is the frequency distribution for sires, where *bl.* = black, *br.* = brown, *b.* = bay, *ch.* = chestnut, *ro.* = roan, *gr.* = grey:

bl.	*bl./br.*	*br./bl.*	*br.*	*br./b.*	*b./br.*	*b.*	*b./ch.*	*ch./b.*	*ch.*	*ch./ro.*	*ro./ch.*	*ro.*	*ro./gr.*	*gr./ro.*	*gr.*
7	4	1	209	0	19	691	0	0	362	1	0	0	0	0	6

Now Mr Yule arranges this in 11 classes:

7	5	209	19	691	0	362	1	0	0	6

and also in three classes, presumably as:

0	—	221	—	710	—	369	—	0	—	0

He then obtains sensibly the same values for the two groupings and speaks of the equality of the pseudo-correlation of ranks thus obtained as marking in some way a limit to the correlation of the variates! Naturally he would obtain almost identical values, because the whole calculation of products and moments turns on the three dominating groups of brown, bay and chestnut and practically all he has done in his arrangements of three groups and eleven groups is to call his sub-range unity in one case and two in the other!

There is, we believe, only one classification possible of these tables on the reasonable assumption that the amount of pigment forms a continuous variate†; namely that which makes a 3×3-fold division between brown, bay and chestnut

* *Phil. Trans.* Vol. 195 A, pp. 122 *et seq.*

† I still see no error in my original classification by amount of pigment; there are more melanin pigment granules in the brown hairs than in the bay, and,—if we disregard the black chestnuts, which are very rare among thoroughbreds as compared with hackneys,—more in the bay than in the chestnut; the latter colour depends more, and in some cases almost entirely, on diffused pigment. K. P.

and either excludes the few roans and greys or, as done below, throws them into the lightest group. The threefold tables are then:

Sire.

Colt.	79	132	47	258
	105	426	132	663
	37	152	190	379
	221	710	369	1300

Sire.

Filly.	58	102	27	187
	66	406	99	571
	25	121	146	292
	149	629	272	1050

Dam.

Colt.	82	73	23	178
	101	319	99	519
	33	129	141	303
	216	521	263	1000

Dam.

Filly.	63	83	23	169
	106	327	88	521
	34	118	158	310
	203	528	269	1000

And the corrected contingency coefficients are:

Sire and Colt: ·41, Dam and Colt: ·45,

Sire and Filly: ·47, Dam and Filly: ·46,

Mean: ·45.

Mr Yule states that the correlation is in this case of the order ·33. Are we again to assert that the mean of the contingency coefficients shows an error of ·12, or is 36 °/₀ in excess of the true correlation, when as we have seen the error we find in even extremely skew distributions is of the order of 4 °/₀ and under?

Mr Yule says that the correlation of his pseudo-ranks is ·33. Well and good, then the correlation for variates would be about ·35, probably more for skew variation, and the factor for correction of class-indices about ·80, or the true correlation is of the order ·35/·80 = ·44, which brings it strangely nearer to the value found by contingency than to Mr Yule's $\frac{1}{3}$!

It is quite true that the values obtained in the original memoir were higher than ·45, but the theory of contingency was not then developed; when a fourfold table had to be formed there were very good reasons for dividing it in the way actually selected. In the first place the division between Chestnut and Bay was physiologically more reasonable than one between Brown and Bay, which would throw the Chestnuts into the Bays. In the next place it was the division nearest to the median and so liable to the least error from either random sampling or skewness. The other, the asymmetrical, divisions are less reasonable because they give one quadrant with only 2 to 3 °/₀ of the total frequency in it, they divide parent and offspring differently, and mix in one or other case Bays with Chestnuts. Let us suppose with Mr Yule that these tables did show a correlation of $\frac{1}{3}$ (which they certainly do not), then we fail to grasp why Mr Yule should not get his Mendelian $\frac{1}{3}$ quite directly without elaborating an erroneous theory of pseudo-ranks and using 3×3-fold and 11×11-fold tables to show approach to a limit.

The Mendelian $\frac{1}{3}$ should come at once by the simple division of the tables into Chestnut and not-Chestnut and it comes pretty closely indeed by taking only the Chestnut groups of the 3×3-fold tables, as indeed it must do if we remember the rarity with which a Chestnut × Chestnut gives other colours. Our four tables become:

Sire.

Colt.	N.-C.	C.	Totals
N.-C.	742	179	921
C. ...	189	190	379
Totals	931	369	1300

Sire.

Filly.	N.-C.	C.	Totals
N.-C.	632	126	758
C. ...	146	146	292
Totals	778	272	1050

Dam.

Colt.	N.-C.	C.	Totals
N.-C.	575	122	697
C. ...	162	141	303
Totals	737	263	1000

Dam.

Filly.	N.-C.	C.	Totals
N.-C.	579	111	690
C. ...	152	158	310
Totals	731	269	1000

The pseudo-rank correlations, ϕ, are:

Sire and Colt: ·3094
Sire and Filly: ·3414
Dam and Colt: ·3030
Dam and Filly: ·3638

Mean = ·3294

Could a better demonstration of the Mendelian $\frac{1}{3}$ correlation be possible?

Now let us look at a similar arrangement of other data in which the true correlation is actually known. We take the following cases with approximately similar total frequencies:

From Table XV.

Stature of Father.

Stature of Son.	*1—4*	*5—8*	Totals
1—4	659	97	756
5—8	143	101	244
Totals	802	198	1000

Actual Correlation ·518.
Yulean ϕ ... ·308.

From Table XIV.

Gaussian Surface for ·5.

	1—4	*5—8*	Totals
1—4	658	98	756
5—8	144	100	244
Totals	802	198	1000

Actual Correlation ·500.
Yulean ϕ ... ·302.

From Table XIII.

Gaussian Surface for ·3.

	1—4	*5—8*	Totals
1—4	634·97	121·03	756
5—8	167·03	76·97	244
Totals	802	198	1000

Actual Correlation ·300.
Yulean ϕ ... ·167.

In other words for any surface approaching the Gaussian the Yulean ϕ for a fourfold table with such frequencies in the total columns must be raised by ·20 in the scale of correlation. Let us consider the same sort of tables in non-Gaussian material, material of Mr Yule's own choosing; we have the following fourfold tables:

From Table XIX.

Age of Wife.

Age of Husband.	*a*	*b—f*	Totals
a	1,135,815	612,932	1,748,747
b—f	110,553	3,458,220	3,568,773
Totals	1,246,368	4,071,152	5,317,520

Actual Correlation ·925.
Yulean ϕ ... ·686.

From Table XVI. *Heredity of Mendelian Couplets.*

Father.

Offspring.	*0*	*1—4*	Totals
0	625	671	1296
1—4	671	2129	2800
Totals	1296	2800	4096

Actual Correlation ·333.
Yulean ϕ ... ·243.

Father.

Offspring.	*0—1*	*2—4*	Totals
0—1	2425	599	3024
2—4	599	473	1072
Totals	3024	1072	4096

Actual Correlation ·333.
Yulean ϕ ... ·243.

From Table XVIII. *Barometer Heights.*

Southampton.

Laudale.	*1—4*	*5—8*	Totals
1—4	2053·25	222·75	2276
5—8	244·25	401·75	646
Totals	2297·5	624·5	2922

Actual Correlation ·780.
Yulean ϕ ... ·531.

Southampton.

Laudale	*1—2*	*3—8*	Totals
1—2	659	321·5	980·5
3—8	233	1708·5	1941·5
Totals	892	2030	2922

Actual Correlation ·780.
Yulean ϕ ... ·566.

From Table XVII. *Ivy Leaf Length.*

First Leaf.

Second Leaf.	Under *6·95*	Over *6·95*	Totals
Under *6·95*	3358	3626	6984
Over *6·95*	3626	49390	53016
Totals	6984	53016	60000

Actual Correlation ·567.
Yulean ϕ ... ·412.

*Whorls of Woodruff**.

Members of First Whorl.

Members of Second Whorl.	Under *8*	*8* and over	Totals
Under *8*	7268	2281	9549
8 and over	2281	1400	3681
Totals	9549	3681	13230

Actual Correlation ·173.
Yulean ϕ ... ·141.

Now let us put together these results in a Table:

Actual Correlation	Yulean for Fourfold	$r-\phi$	Percentage Increase on ϕ
·925	·686	·239	35
·780	·566	·214	38
·780	·531	·249	47
·567	·412	·155	38
·518	·308	·210	68
·500	·302	·198	66
·333	·243	·090	37
·300	·167	·133	80
·173	·141	·032	23

Now these results bring out the important point that whether the distribution of the frequency be Gaussian or not, we may have to add anything from 23 % up to 80 %† to the value of ϕ as found from a fourfold table to obtain the true correlation. For values about ·33 we may have to add anything from 37 % to 80 %. For the tables for coat-colour in horses we must add to the Yulean at a moderate estimate something like 40 %, which gives a value, not near the ·33 of Mr Yule, but near our ·46 and close to the value found by contingency or closer to the values originally assigned by the tetrachoric r_t method.

If Mr Yule continues to assert that the true value of the correlation of these tables is $\frac{1}{3}$, then he may as well argue from the ϕ of the fourfold tables obtained

* Selected as a case of irregular and skew correlation of a low product-moment value; see *Phil. Trans.* Vol. 197 A, p. 325.

† One Ivy Leaf distribution as below actually gave Yulean $\phi = {}$·278 against actual correlation ·567, requiring a 104 % increase on the Yulean!

First Leaf.

Second Leaf.	Under *14·95*	Over *14·95*	Totals
Under *14·95* ...	49562	4054	53616
Over *14·95* ...	4054	2330	6384
Totals ...	53616	6384	60000

by taking as classes Chestnut and non-Chestnut. But in that case his only logical standpoint is to assert that inside the Chestnut and outside the Chestnut there are no hereditary differences of pigment, that Chestnut is a unit character that differs by a unit from all other shades in Bay and Brown. This, however, is not a fact as a very little microscopic examination of horse hair would show him. Pearson's tables themselves indicate that excluding Chestnut, there is correlation of intensity of pigmentation between parent and offspring; further unpublished material indicates that within the Chestnut and for different shades of it the like relation holds. If there be a correlation between parent and offspring inside and outside Chestnut the value for the fourfold tables in the Chestnut and non-Chestnut unit classes must be a *minimum* and not a maximum value as Mr Yule asserts, and his criticism collapses with the fallacies on which he has constructed it.

The chief of these fallacies is the principle that stress of some kind can be laid on the Yulean, or pseudo-rank coefficient proceeding to a limit; the fact is that it can be made anything we please by a suitable choice of divisions. The divisions which for a given number of cells give it a maximum value are those which make the sub-range frequencies of the two variates equal. Every deviation from this equality lessens the value of the Yulean, whether the deviation consists in heaping up the frequency at one or both ends or in the middle of the variate range.

In the following table are a few results for the pseudo-ranks coefficient for 3×3-fold divisions of the Husband and Wife Table (see our p. 224). They fully substantiate the view that it can be made to take almost any value by a proper choice of grouping in the scale classes:

Wife Groups	Husband Groups	Yulean Coefficient
1 : 2 : 3—8	1—6 : 7 : 8	·184
1—6 : 7 : 8	1 : 2 : 3—8	·254
1—3 : 4 : 5—8	1 : 2—7 : 8	·268
1 : 2—7 : 8	1—2 : 3—4 : 5—8	·297
1 : 2—7 : 8	1 : 2—7 : 8	·551
1—4 : 5 : 6—8	1—4 : 5 : 6—8	·605
1—2 : 3—4 : 5—8	1—2 : 3—4 : 5—8	·809

The whole fallacy becomes at once obvious in the light of the class-index correlation correction, for even with a 6×6 or 7×7 table wide changes may be made owing to changes in the classification influencing the class-index correlation*.

Mr Yule has selected the following Series A for Brother-Brother's eye-colour in order to show that ϕ decreases to a limit of less than ·28. But why should he not have taken Series B in order to show that it increases to a value *greater than* ·28?

* Cf. Table XI, p. 218.

Mr Yule accentuating Series A writes (*loc. cit.* p. 619) "The result emphasises the entire non-normality of the eye-colour table. For the normal distribution

Order of Table	Series *A* Divisions	Yulean	Series *B* Divisions	Yulean
2×2	1+2+3 : 4+5+6+7+8	·34	1+2+3+4+5+6 : 7+8	·16
3×3	1+2 : 3 : 4+5+6+7+8	·36	1+2+3+4+5+6 : 7 : 8	·17
4×4	1+2 : 3 : 4 : 5+6+7+8	·33	1+2+3+4 : 5+6 : 7 : 8	·20
5×5	1+2 : 3 : 4 : 5+6 : 7+8	·30	1+2+3+4 : 5 : 6 : 7 : 8	·21
6×6	1 : 2 : 3 : 4 : 5+6 : 7+8	·30	1+2+3 : 4 : 5 : 6 : 7 : 8	·25
7×7	1 : 2 : 3 : 4 : 5+6 : 7 : 8	·29	1+2 : 3 : 4 : 5 : 6 : 7 : 8	·28
8×8	1 : 2 : 3 : 4 : 5 : 6 : 7 : 8	·28	1 : 2 : 3 : 4 : 5 : 6 : 7 : 8	·28

the correlation gradually increases towards the known true value as the number of arrays is increased: with five or eight arrays, notwithstanding the extreme irregularity of the grouping, we have the same moderately good approximation to the correlation as is given by the coefficient of contingency in this case*. For the eye-colour table the correlation *decreases* as the number of arrays is increased."

The fact is we can select series of table divisions for some of which the Yuleans go up, for others they go down, and for still others first go up and then go down. That any series must ultimately reach ·28 is obvious, because that is the value of the only possible 8×8 table, but a 9×9 table might equally well show ·24, and then we suppose that taking the down series Mr Yule would have asserted the limit to be ·24. But suppose Mr Yule's data had stopped at a 5×5 table, then according to the nature of that 5×5 classification, Mr Yule might have found ·21, or ·33† as his limit, for all his tables of lower order must have gone up or down to those limits!

Mr Yule's method if indefinitely continued would lead him to the correlation of ranks, but what relation this would have to the correlation of variates in cases of markedly skew variation with emphasised deviation from linearity no one at present is in a position to say. Some light, however, can be thrown on it by considering the actual deviations produced in calculating means by aid of Mr Yule's hypothesis which places unit distance between each group of individuals, and ultimately between each individual‡. Let d be the range Mr Yule assumes between each individual of a population n and $R=(n-1)d$ the total range. Then

* In this case, namely the Gaussian for ·3 correlation, the Yulean gives ·26 for both 5×5 and 8×8 tables, the corrected contingency gives ·31 and ·30 for these tables; this is Mr Yule's idea of the "*same* moderately good approximation"!

† Using the division $1:2:3:4+5+6+7:8$.

‡ It is remarkable at this stage of statistical progress to find any one so incapable of appreciating Galton's work on first and second prizes as to replace the equal areas occupied by individuals by equal ranges!

it is obvious that Mr Yule's mean character will always coincide with his median, or his

M = mean = character of individual with minimum value of variate $+\frac{1}{2}R$.

$$\sigma^2 = (\text{standard deviation})^2 = \frac{1}{12}\left(1 + \frac{2}{n-1}\right)R^2,$$

or, if n be at all large,

$$\sigma = R/(2\sqrt{3}).$$

It will be clear that for large numbers

$$\sigma = \frac{\text{maximum value} - \text{minimum value of variate}}{2\sqrt{3}},$$

a relation entirely opposed to the practical independence of range and standard deviation in variates with which we are familiar*. But another difficulty at once arises; in actual practice the subranges have all sorts of different values, and we may know one or more of them. Which one of these subranges is to be taken as the standard unit and have the range expressed in terms of it? Some numerical illustrations will emphasise the extraordinary difficulties, not to say contradictions, of Mr Yule's process of treating subranges as equal units in determining correlation.

For example, the Registrar-General gives ages of Husband and Wife from 15 to about 100. Hence by Mr Yule's method:

Mean Age of Husband	= 57·5 years,
Mean Age of Wife	= 57·5 years,
Standard Deviation, Husband	= 18·7639 years,
Standard Deviation, Wife	= 18·7639 years.

The actual values are:

Mean Age of Husband	= 42·8306 years,
Mean Age of Wife	= 40·5838 years,
Standard Deviation, Husband	= 13·0649 years,
Standard Deviation, Wife	= 12·6813 years.

Thus the Yulean values may differ by 30 °/。 to 50 °/。 from the true values.

To assume that the skew distribution is Gaussian will give much better results than this. Let us illustrate it on the very skew Husband and Wife, Barometer, and Ivy Leaf data. In dealing with actual data, we have to express the means of a variety of arrays in terms of a known subrange common to them all, e.g., bay colour in horses or hazel eyes in men. We will apply in succession Mr Yule's hypothesis and the normal curve to the above data.

* For such frequencies as occur in practice it is much safer to take the range about 6 times rather than 3·5 times the standard deviation.

Ages of Husbands in the Eight Groups of Table XIX.
True Mean Age of Husband 42·8306.

Found from Group	By Gaussian		By Yulean	
	Value	Deviation	Value	Deviation
2	38·57	−4·26	46·32	+ 3·49
3	41·14	−1·69	46·32	+ 3·49
4	40·48	−2·35	45·66	+ 2·83
5	40·20	−2·63	45·66	+ 2·83
6	40·13	−2·70	45·66	+ 2·83
7	41·91	−0·92	31·32	−11·51
Mean ...	—	2·43	—	4·49
From whole Range	—	—	57·50	+14·67

In five cases out of six the Gaussian gives better results than the Yulean in the simple matter of finding means even for such a skew distribution as those of ages of Husband and Wife.

Now let us turn to the Ivy Leaves (Table of Breadth) * :

Groups	By Gaussian		By Yulean	
	Value	Deviation	Value	Deviation
5·95— 7·95	11·83	−1·39	14·18	+0·96
7·95— 8·95	12·93	−0·29	11·06	−2·15
8·95— 9·95	12·60	−0·62	11·06	−2·15
9·95—10·95	12·57	−0·65	11·06	−2·15
10·95—12·95	12·87	−0·34	11·18	−2·04
Mean	—	0·66	—	1·89
From whole Range	—	—	19·95	+6·74

It will be seen that with one exception the true mean breadth (13·2148 for this very skew distribution of ivy leaves) would be substantially better found by using the Gaussian than by Mr Yule's assumptions.

Taking the barometric height at Southampton† we have, noting that the actual mean height is 29·9814, the results given in the table on the following page.

In every case the Gaussian, we see, gives markedly better results than Mr Yule's method.

We think it safe to conclude that the Gaussian can be used to give quite a good approach to the means of variates classed in "broad categories"; it is far

* *Phil. Trans.* Vol. 197 A, p. 352. Breadth instead of Length taken for that character appears still more skew.

† *Phil. Trans.* Vol. 190 A, p. 428.

more adequate than the Yulean pseudo-ranks method. It must give better results than a process which concentrates at unit-distances and renders any attempt at class-index correction impossible.

Groups	By Gaussian		By Yulean	
	Value	Deviation	Value	Deviation
29·5	30·036	+ ·055	29·695	− ·287
29·6 to *29·9*	30·006	+ ·025	30·128	+ ·147
30·0 to *30·1*	29·998	+ ·016	30·039	+ ·058
30·2	29·997	+ ·016	30·095	+ ·113
30·3 to *30·5*	30·007	+ ·026	29·784	− ·198
Mean	—	·028	—	·161
From whole Range	—	—	29·750	− ·231

It may be said that Mr Yule has not used the Yulean to find means; in appearance perhaps not; in actuality he certainly has, for all product-moment processes reduce in actuality to finding the means of arrays. In fact:

$$r = \frac{S(n_{xy}xy)/N - \bar{x}\bar{y}}{\sigma_x \sigma_y}$$
$$= \frac{S(n_x x \bar{y}_x)/N - \bar{x}\bar{y}}{\sigma_x \sigma_y}.$$

Here x in the summation term should be really the mean x of the individuals in the class n_x. Thus in finding r we actually use the means $\bar{x}$ and $\bar{y}$ of the two variates, the means $\bar{y}_x$ of the y variate for all individuals in the class n_x of x's, and the mean x of all the individuals in the class n_x. Mr Yule's method, as we have just seen, must lead to big errors in all these means; it also as we have seen leads to big errors in σ_x and σ_y. Hence if r comes out near the true value, this can only arise from a compensation of errors, the exact measure of which has so far only been determined for the case of a Gaussian distribution.

(11) *The Eye-Colour Data.*

We now turn to the eye-colour data for parent and offspring where we think Mr Yule has been led into precisely the same fallacies by his method of pseudo-ranks as in the coat-colour of horses. The colour shades recognised by Francis Galton in his inquiries were:

1. Light Blue.
2. Blue, Dark Blue.
3. Grey, Blue-Green.
4. Dark Grey, Hazel.
5. Light Brown.
6. Brown.
7. Dark Brown.
8. Very Dark Brown, Black.

The book of data presented by Francis Galton to Pearson in 1899 contains under these entries a record of each family, grandparents, parents, offspring, uncles and aunts. No other data concerning the family were provided;—only quite

recently have the original records come into possession of the Eugenics Laboratory and we now find they contain the important desideratum of *age*. It is proposed shortly to revise these tables of eye-colour, paying regard to the increasing pigmentation of the eye in extreme infancy and the decreasing pigmentation in extreme age. The inclusion of all cases, as no age was provided, has possibly something to do with the admitted irregularity of the tables. But does this irregularity invalidate the main conclusion drawn from the tables, i.e. that eye-colour pigment is inherited at the same rate as the measurable physical characters with a correlation lying somewhere between ·46 and ·50? The divisions made originally in the tables were selected purposely and with very definite ends, namely, (i) to give as good a physiological difference between the two groups as possible, (ii) to put into the same class the same eyes of both parents and offspring, and (iii) to get the least probable error by taking the divisions as near the median as was practically realisable. Mr Yule disregards the importance of (ii) and (iii) and gives the values of the tetrachoric r_t found from a number of divisions, several of which have so few individuals in the quadrant that they are extremely untrustworthy. At the time the data were dealt with there was very little available knowledge as to the distinction between a blue and grey eye. They were put into the *same* class, because it was considered that the total pigmentation of the iris of the grey eye was more akin to that of the blue eye, than to that of eyes like hazel with some macroscopic anterior pigment. The difference between the blue and the grey eye was considered to be one of structure rather than of pigment. With what we know of eyes now, we are not prepared to accept the Mendelian classification of eyes into those without and those with anterior pigmentation of the iris. If such a classification were absolutely legitimate, then grey eyes ought to be put with non-blue or with blue according to whether they possess such pigment. On such a classification as we have indicated the ϕ for inheritance of eye-colour between parent and offspring ought to be $\frac{1}{3}$. Accordingly to test the position for the grey eye we ought to consider whether the correlations come out more nearly $\frac{1}{3}$ with the greys put with the browns, or put with the blues, i.e. whether ϕ for a fourfold table is nearer $\frac{1}{3}$ for the division $1+2:3+4+5+6+7+8$ or for $1+2+3:4+5+6+7+8$. There cannot be a moment's hesitation as to which is the more strictly Mendelian division. We find:

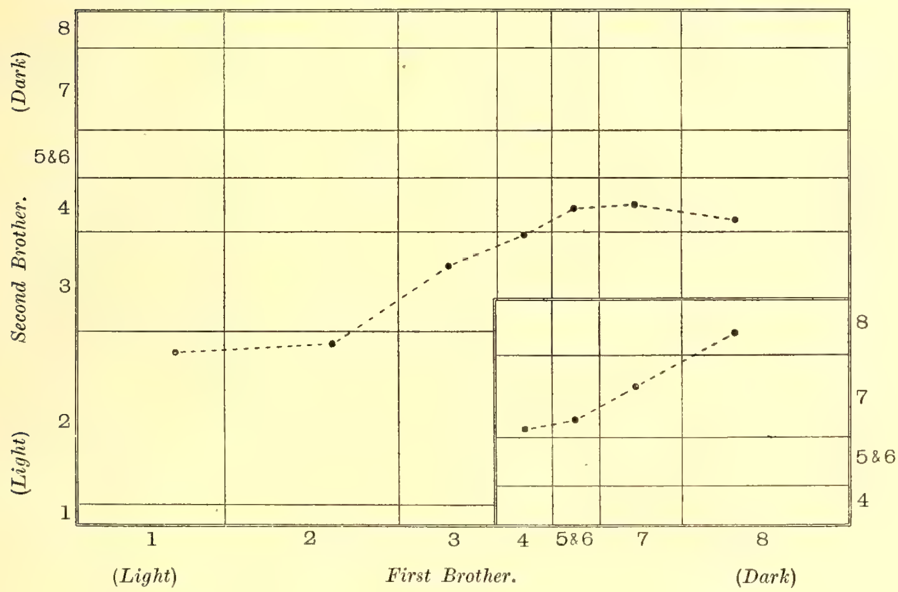
Values of Boas-Yulean ϕ.

Pair	Blue only, i.e. $1+2$	Blue and Grey, i.e. $1+2+3$
Father and Son	·33	·37
Father and Daughter ...	·22	·28
Mother and Son	·28	·32
Mother and Daughter ...	·24	·34
Mean	·27	·33

There seems no doubt that from the Mendelian standpoint, if anterior pigment be taken as a "unit," then pure grey eyes should be included with the blues and the correlation then comes out the true Mendelian third. Thus the original division of the tables between grey and the hazel groups, which appeared at the time the most reasonable physiologically and statistically, is amply justified by the theory of posterior and anterior pigment.

Now Mr Yule tells us on the basis of his erroneous theory of pseudo-ranks that "the average estimated correlation" of these eye-colour "tables is something like $\frac{1}{3}$ not $\frac{1}{2}$" (*loc. cit.* p. 620). We have indicated earlier in this paper (p. 232), that if a fourfold table for continuous variates—skew or Gaussian—give by ϕ a

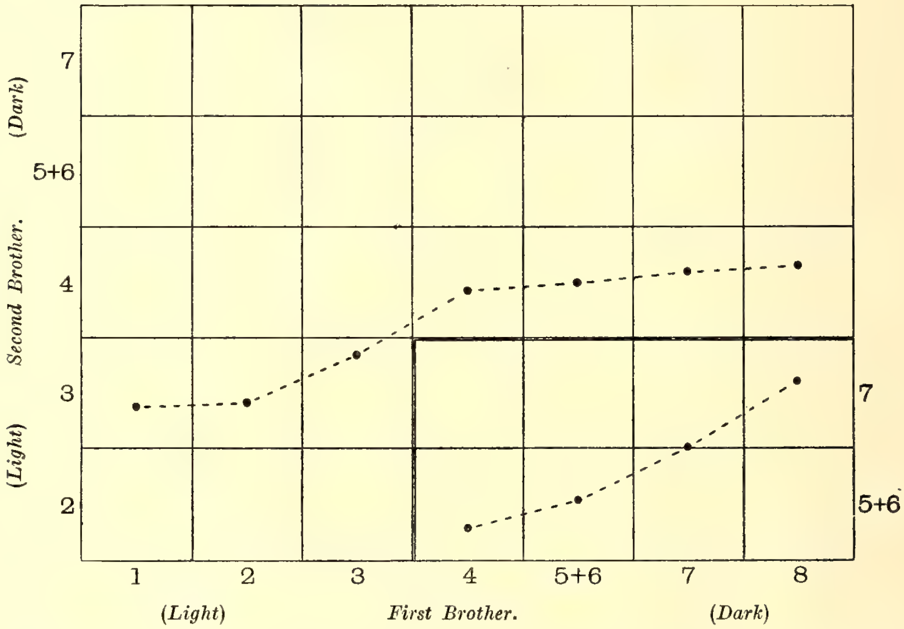
DIAGRAM VI. Regression of eye-colour with eye-colour in brothers, the grade intervals being assumed to give a normal distribution. Inset, the same, colours 4 to 8 only, the grade intervals being of same value as before, to show resemblance within the darker grades only.



correlation of ·33 this must be increased by something like 40 %, if we wish to find the true correlation of the variates. Had Mr Yule asserted that eye-pigmentation was not a continuous variate, that all brown and hazel eyes were alike and differed from blue and grey by the possession of some mysterious unit character, then he would have been justified by the mere fourfold tables in asserting that the correlation was $\frac{1}{3}$. In doing this he would have taken up the simple Mendelian position. That he did not do so was probably owing to the fact that he recognised, as the tables indeed show, that within the two groups, $1 + 2 + 3$ the blue-grey group and $4 + 5 + 6 + 7 + 8$ the hazel-brown group, there was distinct heredity of sub-divisions. The father with hazel eyes has an offspring less pigmented than the father with dark brown eyes. That such heredity exists must dispose at once of a

correlation of $\frac{1}{3}$ as representing the heredity of eye-colour in man. And here it may be well to point out a very distinct difference between the problem which Pearson set himself and the problem for which the Mendelians profess to provide a solution—both speak of the heredity of eye-colour—but in no sense deal with the same problem*. The Mendelian says: "I treat the heredity of the character, presence or absence of anterior pigment." The correlation in this case should be $\frac{1}{3}$. This problem, however, is not that of the heredity of the various grades of pigment in the iris. There is very little doubt that grade of pigment is a perfectly continuous variate. The pigment of the iris is not solely confined to anterior and posterior faces, and quite different grades of pigment can occur in both these

DIAGRAM VII. Regression of eye-colour of brothers, the grade intervals being supposed equal. Inset, the same, colours 4 to 8 only, to illustrate the reduction in the regression—which is here the correlation—produced on the pseudo-rank hypothesis.



situations, until you get in the albinotic eye almost a complete absence of pigment. The accompanying diagrams, which have been drawn by two different methods, show how the heredity of eye pigment holds as well inside the blue-grey group as inside the hazel-brown group. They have been obtained in two ways, (i) by the use of the Gaussian to obtain the means of the arrays, and the relative ranges of

* Even in the treatment of this problem we have only the papers of Hurst and Davenport which have been assumed to confirm each other. As a matter of fact on Hurst's postulates as to the methods of observation, Davenport ought to have reached results discordant with Mendelism instead of confirmatory! For the nature of these authors' work on similar problems, see p. 209 above and *Biometrika*, Vol. VII. p. 403 and Vol. VIII. pp. 269, 271, 272.

the eye-colour, (ii) by Mr Yule's method of unit groupings to indicate that even that method shows the same result, the inheritance of the intensity of pigment inside the two main groups*. In an inset figure we have given, by the same methods, the increasing pigmentation of the second brother as the pigmentation of the first brother increases, when we remove all the blue-grey group from consideration. The general weakening of the correlation produced by using the method of pseudo-ranks will be obvious if Diagrams VI and VII are compared. The later diagrams, Nos. VIII and IX, show precisely the same point for the parental eye-colour data. It will be clear that if $\frac{1}{3}$ were the limit to the correlation of these eye-colour tables, then all the correlation within the blue-grey and hazel-brown groups ought to have disappeared. Instead of this we find that the $\frac{1}{3}$ is only the correlation on the Mendelian hypothesis that there is a unit difference between individuals in the one group and individuals in the other; whatever this "unit difference" may refer to, it does not refer to a quantitative difference in pigmentation, because there is correlation within the groups.

We have then the following results for fourfold parental eye-colour tables, the division being made between marked and less marked anterior pigment:

Classes	Variates supposed to be discrete. Yulean ϕ	Variates supposed to be continuous and Gaussian. Tetrachoric r_t
Father and Son	·37	·55
Father and Daughter ...	·28	·44
Mother and Son	·32	·48
Mother and Daughter ...	·34	·51
Mean	·33	·49

Mr Yule's statement that the correlation is "something like $\frac{1}{3}$, not $\frac{1}{2}$" amounts to the denial that the variate is continuous; and the slightest inspection of our diagrams shows that the variate is continuous; the question therefore turns on whether the Gaussian assumption gives a reasonable approximation to the influence of this continuity in increasing the correlation. We have shown that the fourfold table result (ϕ) obtained by treating the variate as discrete requires, whether the distribution be skew or Gaussian, to be increased by amounts ranging from 37 °/。 to 80 °/。, when the value of ϕ is ·3 or upwards. We have accordingly little hesitation in asserting that the true correlation exceeds $\frac{1}{3}$ by at least 40 °/。. Another way of approaching the problem is the one adopted by Pearson, when he found the want of stability in the tetrachoric r_t as applied to these eye-colour tables, the method of mean square contingency.

* Note how the slope of this regression line is ·5 for the main portion of Diagram VI.

Making the proper corrections we find:

Order of Table	Nature of Classification	Father and Son	Father and Daughter	Mother and Son	Mother and Daughter
7 × 7	1 : 2 : 3 : 4 : 5+6 : 7 : 8	·57	·48	·50	·43
6 × 6	1 : 2 : 3 : 4 : 5+6 : 7+8	·58	·50	·48	·44
5 × 5	1+2 : 3 : 4 : 5+6 : 7+8	·55	·52	·50	·44
1st 4 × 4	1+2 : 3+4 : 5+6 : 7+8	·53	·50	·45	·41
2nd 4 × 4	1+2 : 3 : 4 : 5+6+7+8	·54	·56	·50	·45
1st 3 × 3	1+2 : 3+4 : 5+6+7+8	·52	·45	·45	·41
2nd 3 × 3	1+2+3 : 4 : 5+6+7+8	·60	·51	·56	·50
Mean*	—	·55	·50	·49	·44
Values originally given by Pearson†		·55	·44	·48	·51

That the eye-colour tables present anomalies has been fully admitted, but the average mean square contingency based on 28 groupings of these four tables gives the mean value ·4961; the mean given by Pearson originally from what he considered and we still consider the natural fourfold division was ·4947. We have no experience whatever of the average of a large number of corrected and wholly *a priori* unselected contingencies giving a result too high by 33 °/₀ of its value! On the contrary, as we have shown on pp. 221—226 the contingency method appears on the average to give values slightly in defect of the true correlation.

A study of our diagrams shows that it is in the very small groups 1 and 8 containing only 2·5 °/₀ to 4·5 °/₀ of parents that the marked deviations from linearity of the regression occur, that is to say in the light blue-eyed and very dark brown parents' offspring. We suggest here that a considerable number of the light blue parents may have been erroneously classified, on account of extreme age, and a considerable number of the offspring of very dark-eyed parents may have been classified as light blue because of extreme infancy. Both may really belong to the mediocre classes. It is only in the extreme classes of small frequency that the effect of this shifting of the mediocre would be sensible. The Diagrams IX and X (pp. 248 and 250)‡ show that in the bulk of cases, from group 2 to 6 or even to 7, the regression line has a slope of over ·5; the inset figures show that this regression is maintained inside the brown group when we remove the blues. We have already given similar diagrams for the Brother-Brother table. These again show that in the centre of the figure the slope of the regression line is at least ·5, while the defect at the tails very probably indicates the results of

* Calculated from the corrected contingency to four decimal places.

† *Phil. Trans.* Vol. 195 A, p. 106. The division taken is that which for the use of a single division only is physiologically and statistically the most reasonable; see above, p. 238.

‡ The Diagrams VI and IX were obtained by applying a normal curve to each array to obtain the position of the mean on the assumption that the range of groups 3 and 4 remains the same on the scale.

including cases in which one brother was an infant. The second diagram (p. 240) worked out by Mr Yule's pseudo-ranks indicates also the heredity within the brown group when blues are excluded, but it obviously distorts the whole system and gives the spurious appearance of far less correlation.

Another manner of illustrating the idle character of Mr Yule's assertion that the true parental correlation is about $\frac{1}{3}$ may be obtained from a comparison of the following tables:

A	*1+2*	*3*	*4*	*5+6*	*7+8*	Totals
1+2	162	87	48	16	22	335
3	102	78	52	19	33	284
4	41	37	28	11	20	137
5+6	27	28	22	9	19	105
7+8	26	34	30	14	35	139
Totals	358	264	180	69	129	1000

B	*1+2*	*3*	*4*	*5+6*	*7+8*	Totals
1+2	194	70	41	9	21	335
3	83	124	41	13	23	284
4	25	34	55	11	12	137
5+6	27	12	19	24	23	105
7+8	29	24	24	12	50	139
Totals	358	264	180	69	129	1000

C	*1+2*	*3*	*4*	*5+6*	*7+8*	Totals
1+2	193	84	38	10	10	335
3	100	85	54	19	26	284
4	34	39	31	12	21	137
5+6	18	28	25	11	23	105
7+8	13	28	32	17	49	139
Totals	358	264	180	69	129	1000

Now the changes that will convert *A* into *B* are given by the scheme:

B−A	*1+2*	*3*	*4*	*5+6*	*7+8*
1+2	+32	−17	− 7	− 7	− 1
3	−19	+46	−11	− 6	−10
4	−16	− 3	+27	0	− 8
5+6	0	−16	− 3	+15	+ 4
7+8	+ 3	−10	− 6	− 2	+15

That is to say to obtain *B* from *A* we must take between $\frac{1}{7}$ and $\frac{1}{8}$ of the material of *A* and transfer it to the diagonal cells from those cells away from the diagonal, the only failure of this rule is the $+3$ in the cell $1+2$: $7+8$. There

can be no doubt that the correlation of B is very much greater than the correlation of A. To convert C into A the changes are given by the scheme:

$B-C$	$1+2$	3	4	$5+6$	$7+8$
$1+2$	+1	−14	+3	−1	+11
3	−17	+39	−13	−6	−3
4	−9	−5	+24	−1	−9
$5+6$	+9	−16	−6	+13	0
$7+8$	+16	−4	−8	−5	+1

Here again the candid reader will, we believe, admit that B has more correlation than C. The general difference is again a transfer to the diagonal cells of frequency to the right and left of this diagonal, the total number so transferred is 78; but the transfer is rendered more complex by the appearance of 36 units in the cells in the top right-hand and bottom left-hand corners. This transfer is only *half* the size of the other towards the diagonal, and it probably measures the frequency of senile blues in the parents and of infantile blues in the offspring. We hold that we must still credit B with more correlation than C, although this secondary movement can now be well recognised. B is the contingency table for Father and Son's eye-colour (see p. 186). A is a table for a Gaussian surface with ·3 correlation, C is a similar table for a Gaussian surface with ·5 correlation, both being adjusted to give frequencies to unit places only with the same marginal totals as the Father and Son eye-colour table. These comparisons suffice to indicate that the correlation of Father and Son is far greater than ·30 and is probably slightly greater than ·50.

Precisely similar treatment of the Brother-Brother eye-colour table leads to like conclusions, as the reader may judge from what follows.

Consider the three tables:

A	$1+2$	3	4	$5+6$	$7+8$	Totals
$1+2$	398	283	113	66	71	931
3	283	281	135	87	115	901
4	113	135	72	50	73	443
$5+6$	66	87	50	36	58	297
$7+8$	71	115	73	58	111	428
Totals	931	901	443	297	428	3000

B	$1+2$	3	4	$5+6$	$7+8$	Totals
$1+2$	496	224	63	68	80	931
3	224	418	97	56	106	901
4	63	97	168	47	68	443
$5+6$	68	56	47	70	56	297
$7+8$	80	106	68	56	118	428
Totals	931	901	443	297	428	3000

C	$1+2$	3	4	$5+6$	$7+8$	Totals
$1+2$	471	278	94	48	40	931
3	278	300	140	86	97	901
4	94	140	78	55	76	443
$5+6$	48	86	55	42	66	297
$7+8$	40	97	76	66	149	428
Totals	931	901	443	297	428	3000

Now mark what change must be made in A to correct it to B; this is given by the scheme:

$B-A$	$1+2$	3	4	$5+6$	$7+8$
$1+2$	+98	− 59	− 50	+ 2	+9
3	− 59	+137	− 38	− 31	− 9
4	− 50	− 38	+96	− 3	− 5
$5+6$	+ 2	− 31	− 3	+34	− 2
$7+8$	+ 9	− 9	− 5	− 2	+7

Can there be any doubt that the scheme $B-A$ marks immensely increased correlation? To pass from A to B, we must accumulate 372 individuals along the diagonal as against 11 individuals drawn towards the corners away from this diagonal. We take it that only the most captious person could possibly deny that to reach B from A, we must transfer individuals in a manner which markedly increases the correlation.

Now consider the change which must be made in C to obtain B. It is given by the scheme:

$B-C$	$1+2$	3	4	$5+6$	$7+8$
$1+2$	+25	− 54	− 31	+20	+40
3	− 54	+118	− 43	− 30	+ 9
4	− 31	− 43	+90	− 8	− 8
$5+6$	+20	− 30	− 8	+28	− 10
$7+8$	+40	+ 9	− 8	− 10	− 31

To pass from C to B we must transfer 261 individuals to the diagonal; but there is an outward movement of 69 to each corner and a resulting defect of -31 in the fifth diagonal cell. There is a total movement of 261 individuals towards greater, and of 169 towards lesser correlation. We should have no hesitation in saying that the correlation of B is higher than that of the C table. We would go further and say that until we find a table D in which these movements towards the diagonal and towards the outer corners nearly balance we have not yet reached a table with a correlation equal to that of B. Consider the table:

D	*1+2*	*3*	*4*	*5+6*	*7+8*	Totals
1+2	494	275	88	42	32	931
3	275	308	142	86	90	901
4	88	142	81	56	76	443
5+6	42	86	56	44	69	297
7+8	32	90	76	69	161	428
Totals	931	901	443	297	428	3000

We have the scheme:

B − D	*1+2*	*3*	*4*	*5+6*	*7+8*
1+2	+ 2	− 51	−25	+26	+48
3	−51	+110	−45	−30	+16
4	−25	− 45	+87	− 9	− 8
5+6	+26	− 30	− 9	+26	−13
7+8	+48	+ 16	− 8	−13	−43

Here 225 individuals must be transferred to the diagonal to get from *D* to *B*, i.e. there is an increase to this extent of the correlation, but 223 have got to be transferred outwards to the ends of the other diagonal, although not in such a concentrated fashion. There can be small doubt, we think, that the correlation of *B* is nearly, if not slightly greater than, that of *D*. *B* is the correlation table for Brothers, *A*, *C* and *D* have correlations respectively of ·28, ·45 and ·50. We feel confident in asserting that the unknown correlation of *B* is far nearer to that of *C* or *D*, than of *A*, which Mr Yule gives as its limit! Nor can the reader who examines these tables fail to obtain true insight into the peculiar nature of the correlation of *B*. It is clear that *B* fails of normality because there is an excess of dark brothers having brothers with blue eyes. The excess is greater than in the case of father and son and this is actually what we should expect, if the excess be due to the inclusion in the record of *infants*. For an infant will be reckoned once as a son, but if he is one of n brothers, he will appear $n(n-1)$ times. If our surmise be correct there may be no failure at all of approximately Gaussian frequency in these eye-colour tables, but anomalous lumps in the second and fourth quadrants due to this unconscious inclusion of infants. If these lumps, which the inclusion of a dozen to twenty infants would suffice to produce, be removed, we should expect the correlation to rise as in the Huxley Lecture to about ·60. What the final value would be after the correction for the more gradual changes in eye-colour, which occur after infancy*, remains to be determined; we may be quite certain, however, that it is likely to be 80 to 100 % above Mr Yule's maximum of ·28.

* The influence of age on parental and fraternal tables of eye-colour will be shortly considered *de novo*.

But let us look at the matter from other standpoints. We have the stature data for Father and Son arranged in the eye-colour groups. Let us arrange it in a fourfold table, precisely as we did the eye-colour in order to demonstrate Mendelism, i.e.

Stature of Father.

Stature of Son.	*1+2+3*	*4+5+6+7+8*	Totals
1+2+3 ...	464	155	619
4+5+6+7+8	158	223	381
Totals ...	622	378	1000

We may give these a class-name, say, Short and Tall—corresponding to Short Muzzle and Long Muzzle, or to "Over 30" and "Under 30" egg hens, etc., etc. We have at once a Mendelian table:

	Short Fathers	Tall Fathers	Totals
Short Sons ...	464	155	619
Tall Sons ...	158	223	381
Totals ...	622	378	1000

Now assume a discrete unit between Short and Tall and find the Boas-Yulean ϕ. It is

$$\phi = \cdot 335,$$

and as its Mendelian value should be ·333, the agreement is extraordinary! Tallness and shortness in man are clearly Mendelian presence and absence of a unit character! The eye-colour tables at the same divisions show exactly the same result. Surely Mr Yule is satisfied that the correlation of stature in man is of the order $\frac{1}{3}$? Yet the tetrachoric r_t in these cases at the same division gives for eye-colour ·55 and for stature ·51, i.e. is some 60 to 70 °/ₒ greater. Now let us treat these two cases by an identical process; we use a normal horizontal and a normal vertical scale, and we determine all ranges in terms of the range of the frequency of groups 3 + 4 treated as of length h. We find the mean values of the arrays of sons for each group of fathers by assuming that the means will be approximately given if the array be treated for this purpose as Gaussian (see p. 236). The results are given in the table on p. 249.

Can any reader who examines these data deny the remarkable parallelism of the two cases? Let him also look at the Diagrams VIII and IX (p. 248) drawn for stature and eye-colour and ask himself whether it is possible to make any distinction between the cases of inheritance of eye-colour and inheritance of stature, beyond the irregularity depending on the group of 36 cases (3·6 °/ₒ of the whole) of dark blue-eyed fathers. The two regression lines run with almost complete agreement, the stature data being a little more regular than the eye-colour data.

Diagram VIII. Regression in Stature. Father and Son. Based upon means obtained by Gaussian assumptions from graded but unmeasured stature groups, equal to the eye-colour groups.

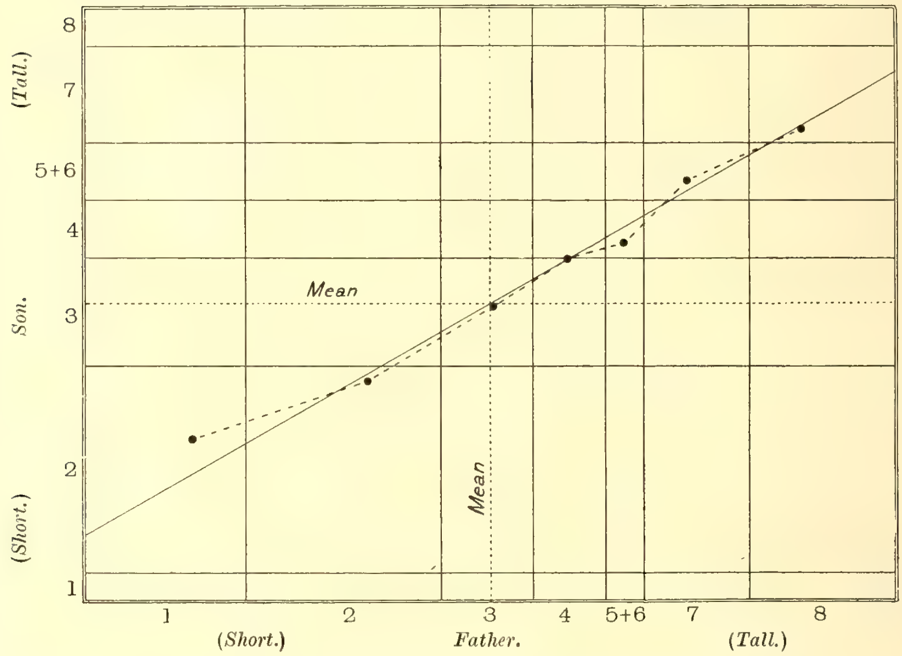
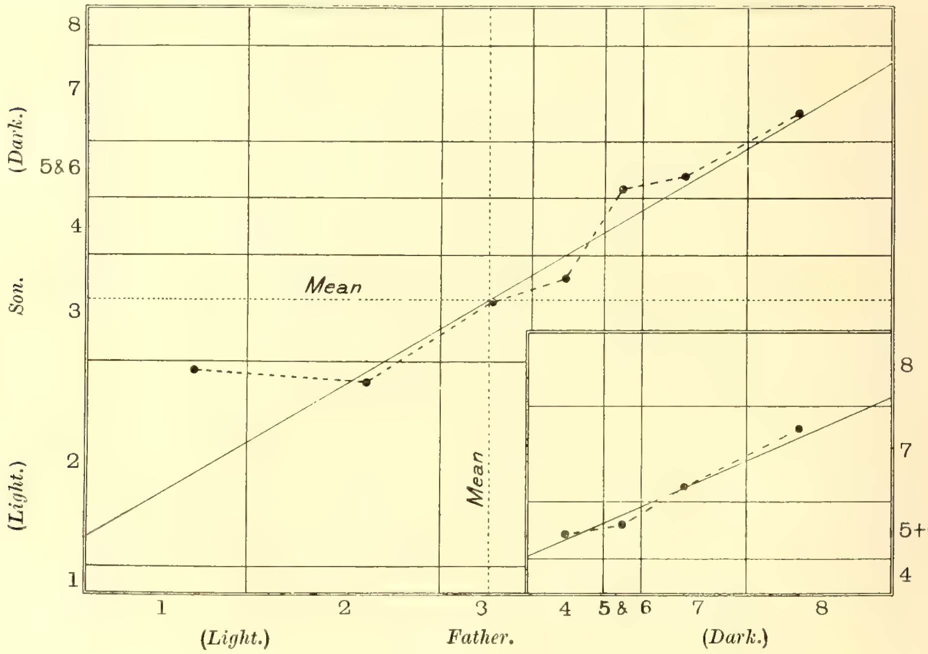


Diagram IX. Regression in Eye-colour. Father and Son. Based upon means obtained by Gaussian assumptions. To be compared with the like regression for stature.



Mean Son	Eye-Colour	Stature
Group *1*	− ·4336h	− ·8255h*
,, *2*	− ·5065h	− ·4778h
,, *3*	− ·0166h	− ·0246h
,, *4*	+ ·1319h	+ ·2686h
,, *5+6*	+ ·6704h	+ ·3611h
,, *7*	+ ·7513h	+ ·7353h
,, *8*	+1·1414h	+1·0455h*
Standard Deviation σ_F	·8247h	·8247h
,, ,, σ_S	·8932h	·8932h
η uncorrected from above	·5217	·5074
η corrected	·5420	·5224
Slope of Regression Line, from η used as r ...	·5870†	·5658
True value of r	?	·5189
Tetrachoric r_t‡	·5503	·5104

Can any one continue to assert with Mr Yule that the correlation for eye-colour is about $\frac{1}{3}$ and that for stature is $\frac{1}{2}$? Is it not clear that for eye-colour in Father and Son the value given by Pearson in 1900 is within ·02 of the true value?

A similar figure (Diagram X) is given on p. 250 for Mother and Son. This is more irregular in the terminal groups (which contain 3·6 °/₀ and 4·6 °/₀ of the frequency only), but it shows the same points. The uncorrected $\eta = $ ·4930, the corrected $\eta = $ ·5109 and the regression is ·5459. These are in quite good agreement with the value of r as found from contingency, i.e. ·4885 (see p. 241), but differ slightly in excess from the value given by the tetrachoric r_t for the "Mendelian" table, i.e. correlation ·4817 and regression ·5145. It is clear from the diagram that r, as regards η, has been lessened by the deviations from linearity in the terminal classes§. Still the slope of the regression line shows that we are far from dealing with a correlation of $\frac{1}{3}$ for there is approximate equality in the variations of the eye-colour in mother and son, $\sigma_M = $ ·8574 h and $\sigma_S = $ ·9161 h.

(12) *The Vaccination Data.*

We now turn to the question of vaccination and recovery from small-pox. Mr Yule apparently considers either his coefficient of association, or his coefficient of colligation, to be the right coefficient to use here. Now both these coefficients are

* The means in the case of the two terminal arrays had to be somewhat differently treated as there is no frequency in the first array of groups 5—8 and in the last array of groups 1 + 2. Accordingly if the h of the table be h_{3+4}, the mean of the first array was found in terms of h_2, i.e. the range of group 2 for that array, and the mean of the last array from h_{5+6+7} for that array, i.e. the range of groups 5 + 6 + 7 for that array. h_2 and h_{5+6+7} were then expressed in terms of h_{3+4}, i.e. of h of the table by means of the relations between these subranges in the marginal total frequencies for sons.

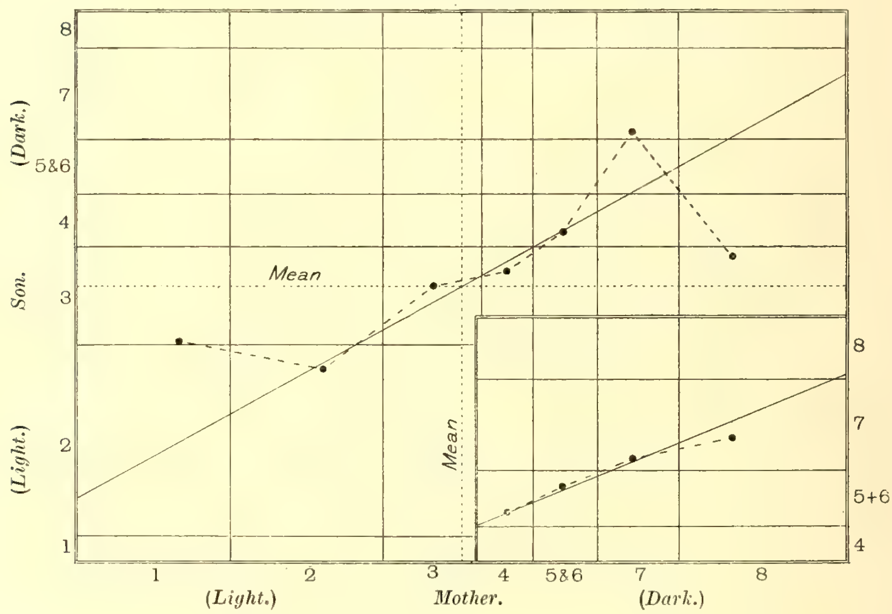
† To four figures the value as given in 1900 was ·5159.

‡ For divisions as originally given in the *Phil. Trans.* memoir, Vol. 195 A, p. 106.

§ See remarks, p. 164 above, on the true measure of correlation.

unchanged by the artificial selection in which Mr Yule sees* "a most important property and one of special importance in such cases as those I have chosen for illustrations," i.e. vaccination and recovery. On the other hand the ϕ, or pseudo-rank correlation, is for a fourfold table immensely influenced, as we have shown, by selection of this artificial kind. If this ϕ, however, be artificially selected, so that it deviates widely from the ϕ of the *data actually provided*, then this artificial value of ϕ—for a table, say, in which the number of vaccinated has been made equal to the number of unvaccinated and the number of deaths equal to the number of recoveries—is Mr Yule's coefficient of colligation. From our standpoint it is hard to conceive a stronger argument against this coefficient of colligation than the fact that it is unequal to ϕ, unless you have artificially

DIAGRAM X. Regression in Eye-colour. Mother and Son. Based upon means obtained by Gaussian assumptions.



doctored ϕ to bring the two into agreement. What is the probable error of ϕ thus doctored or how is it related to the probable error of ϕ found from the undoctored material? Mr Yule, however, writes: "We have therefore the important theorem briefly mentioned without proof in p. 17—*the coefficient of colligation ω for any table is the product-sum correlation r*† *for the equivalent symmetrical table.* These two coefficients r and ω form, accordingly, a natural pair, the first giving the actual correlation in the given table, the second the correlation in a derived table of standard form, thus enabling us to compare the

* *Loc. cit.* p. 587.

† Mr Yule here as elsewhere terms ϕ the "product-sum correlation" and uses for it the letter r. This is wholly unjustifiable, it is merely a correlation of pseudo-ranks and not of true variates at all.

two tables freed from the effects of 'selecting varying proportions of A's and B's'" (*loc. cit.* p. 597). The italics are Mr Yule's. We have rarely come across more specious reasoning. A coefficient is selected which for one type of *artificial* selection is not changed, and this peculiarity is termed a special and important feature of the coefficient. Another coefficient which is intensely subject to this selection is then commended because it can be selected so as to agree with the first; and Mr Yule terms them a "natural pair." Algebraically Mr Yule starts with the table $\frac{a \mid b}{c \mid d}$, and the Q for this is $(ad - bc)/(ad + bc)$. This value of Q remains unchanged if we write the table $\frac{\sqrt{ad} \mid \sqrt{bc}}{\sqrt{bc} \mid \sqrt{ad}}$ where the marginal frequencies have now been rendered artificially all the same. The original ϕ however was

$$(ad - bc)/\sqrt{(a + b)(a + c)(b + d)(c + d)},$$

and has changed to

$$(\sqrt{ad} - \sqrt{bc})/(\sqrt{ad} + \sqrt{bc}),$$

which is Mr Yule's colligation and a function of Q. The original ϕ has many important properties, it is Pearson's mean square contingency for a fourfold table and determines the *probability* of the two variates being independent; it is also the correlation of means if they be measured from their dividing lines in terms of their standard deviations, supposing the material continuous and approximately normal. The new ϕ possesses neither of these really important properties, it is the ϕ of the table after most artificial selection, and that it agrees with Mr Yule's coefficient of colligation neither gives validity to that coefficient, nor endows it with any new property whatever.

Mr Yule has not taken the trouble to see what sort of effect his artificial selection really has on actual material. We propose to illustrate it on certain vaccination data. We give, as Table XXI, the table of Glasgow data published in *Biometrika*, Vol. VII. p. 257. It is not true that the haemorrhagic and confluent cases all die, but a very large percentage of them die and it would be not far from the fact to represent the data by a fourfold table:

	Deaths	Recoveries	Totals
Vaccinated ...	273	1301	1574
Unvaccinated ...	65	50	115
Totals ...	338	1351	1689

This table is only illustrative, not, of course, a rigorous experience.

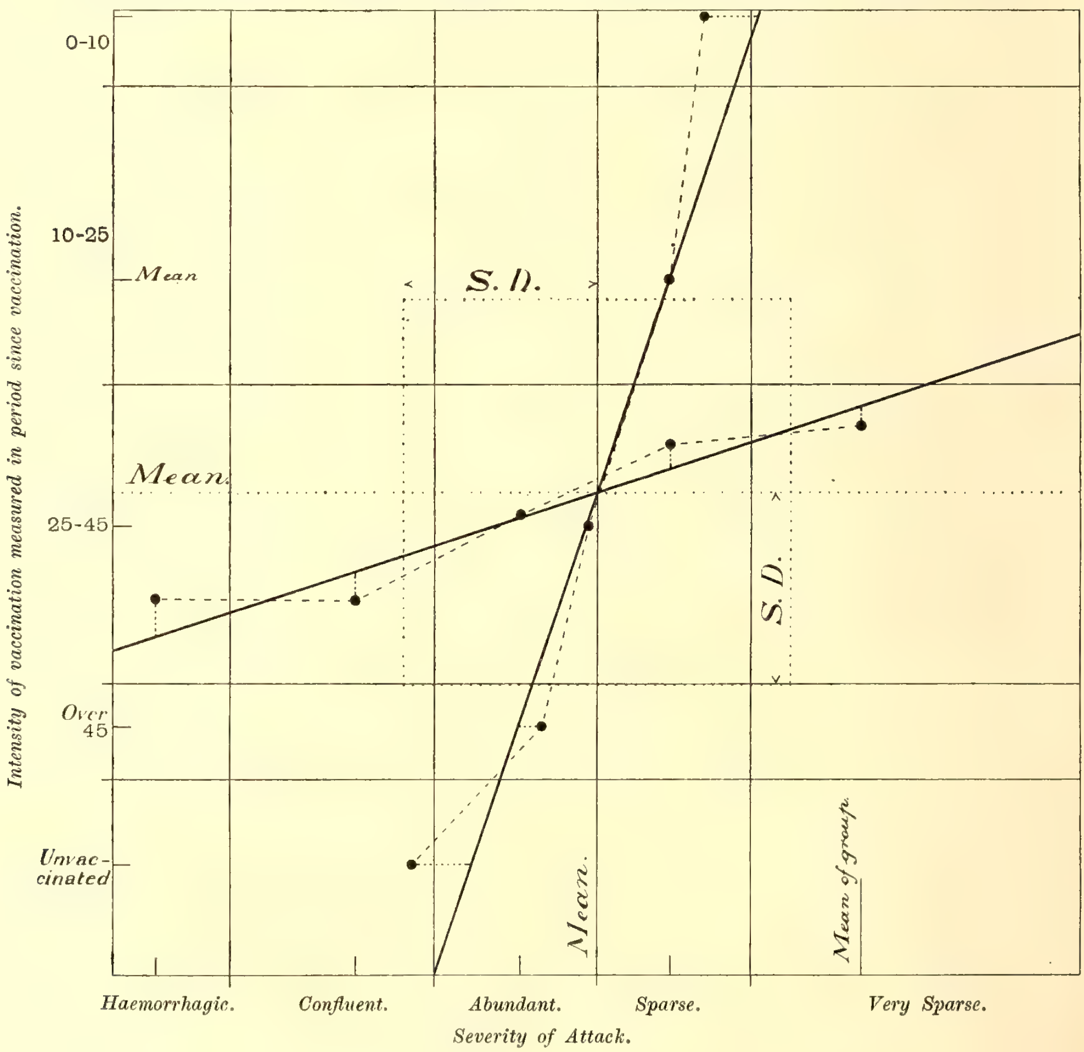
Now we have not assumed Table XXI to be a Gaussian distribution, but we have found the means of the arrays by assuming each to be represented by a Gaussian curve, a process which, as we have seen, gives a fairly close result even for skew material. We have also, in order to get scales of a reasonable character, arranged

TABLE XXI. *Severity of Attack and Strength of Immunity due to Vaccination. City of Glasgow Hospital.*

Severity of Attack.

Grade of Vaccination Immunity.	Haemorrhagic	Confluent	Abundant	Sparse	Very Sparse	Totals
0—10	0	1	6	11	12	30
10—25	5	37	114	165	136	457
25—45	29	155	299	268	181	932
Over *45*	11	35	48	33	28	155
Unvaccinated	4	61	41	7	2	115
Totals ...	49	289	508	484	359	1689

Diagram XI. Regression of the intensity of vaccination upon severity of attack and of severity of attack upon intensity of vaccination, showing the continuity of the relationship. Measured on Gaussian scales.



the severity of attack and immunity due to vaccination on Gaussian bases. This is done in Diagram XI, and the corresponding regression lines are drawn. These were calculated by supposing the range of the mid or 'abundant' group the same for all arrays of immunity, and the group 25—45 the same for all arrays of severity. It will be seen at once from this diagram that the division into vaccinated and unvaccinated is a merely verbal one, that Mr Yule is playing with words, not dealing with the realities under class-names*. Immunity and severity are continuously changing quantities and they vary in almost linear relationship to each other, the greater the intensity of the immunity associated with vaccination the less severe is the attack. We do not think the deviations from linearity of regression found in this case differ sensibly from the sort of values that frequently occur in the case of the regression line found by the product-moment method.

In order, however, that Mr Yule may not attribute this result to the use of the Gaussian distribution to find the means, we applied his method of pseudo-ranks to determine the means of the arrays and then put in the regression lines on his hypothesis that the rank of a huge bracket is a suitable unit in which to measure correlation. This is done in the lines A of Diagrams XII and XIII. Although the previous scheme is much distorted, we see that on Mr Yule's own hypothesis the two variates behind the "vaccination—non-vaccination" and "death—recovery" classifications are continuous, steadily increasing in one direction, and that there is the correlation of a continuous two-variate system behind them.

Now let us proceed to do exactly like what Mr Yule has done, namely make divisions at vaccinated and non-vaccinated, and between confluent and abundant, and dress the table so that there shall be equal numbers of vaccinated and unvaccinated and of the very severe and less severe cases. The table thus transformed is Table XXII. The difficulty now is to know what to do with the marginal scales; if these in the natural condition were normal, they are hardly likely to remain so after selection. But these scales—especially as we have found the correlations by the η-process—are not of first class importance except for exhibition of the results graphically. We have therefore retained the old scales and merely calculated the means of the arrays condensed at the old means of the old scale of years since vaccination (see Diagram XIV). In order that the reader may see how little assumption is thus made we have had a second diagram drawn indicating what

* "The division may also be vague and uncertain: sanity and insanity, sight and blindness, pass into each other by such fine gradations that judgments may differ as to the class in which a given individual should be entered. The possibility of uncertainties of this kind should always be borne in mind in considering statistics of attributes: whatever the nature of the classification, however, natural or artificial, definite or uncertain, the final judgment must be decisive; any one object or individual must be held to possess the given attribute or not." *Theory of Statistics*, pp. 8, 9. Similar words are used by Mr Yule in *Biometrika*, Vol. II. p. 121: "The judgment must however be finally *decisive*; intermediates not being classed as such even when observed." We can hardly conceive statements more liable to prejudice the mind of a wavering recorder of actual data. No wonder Mr Sanger said that "Mr Yule's work would be the work for Mendelians"!

TABLE XXII. *Immunity and Severity of Vaccination Table after a Yulean Equalisation Process.*

Period since Vaccination	Haemorrhagic	Confluent	Abundant	Sparse	Very Sparse	Totals
0—10	0	0·9	2·8	5·1	5·6	14·4 } 844·5
10—25	4·4	32·8	52·8	76·4	63·0	229·4 }
25—45	25·7	137·4	138·5	124·1	83·8	509·5 }
Over *45*	9·7	31·0	22·2	15·3	13·0	91·2 }
Unvaccinated	37·1	565·5	198·4	33·9	9·6	844·5
Totals	76·9	767·6	414·7	254·8	175·0	1689·0

$\underbrace{\text{Haemorrhagic} + \text{Confluent}}_{844\cdot5}$ $\underbrace{\text{Abundant} + \text{Sparse} + \text{Very Sparse}}_{844\cdot5}$

DIAGRAMS XII AND XIII. Intensity of Vaccination and Severity of Disease treated by the method of pseudo-ranks.

Assumed unit grades. (*A*) original data, (*B*) Haemorrhagic and Confluent made 50 °/₀ and vaccinated made 50 °/₀.

(i) Regression of intensity upon severity. (ii) Regression of severity upon intensity.

DIAGRAM XII.

(i)

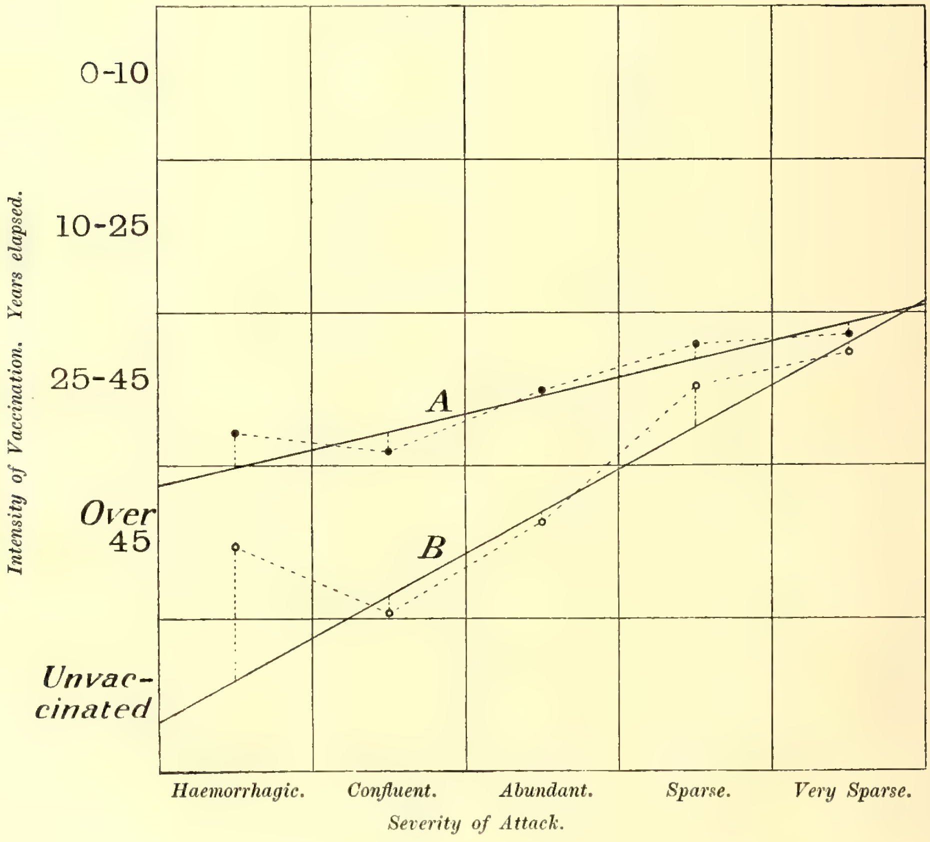
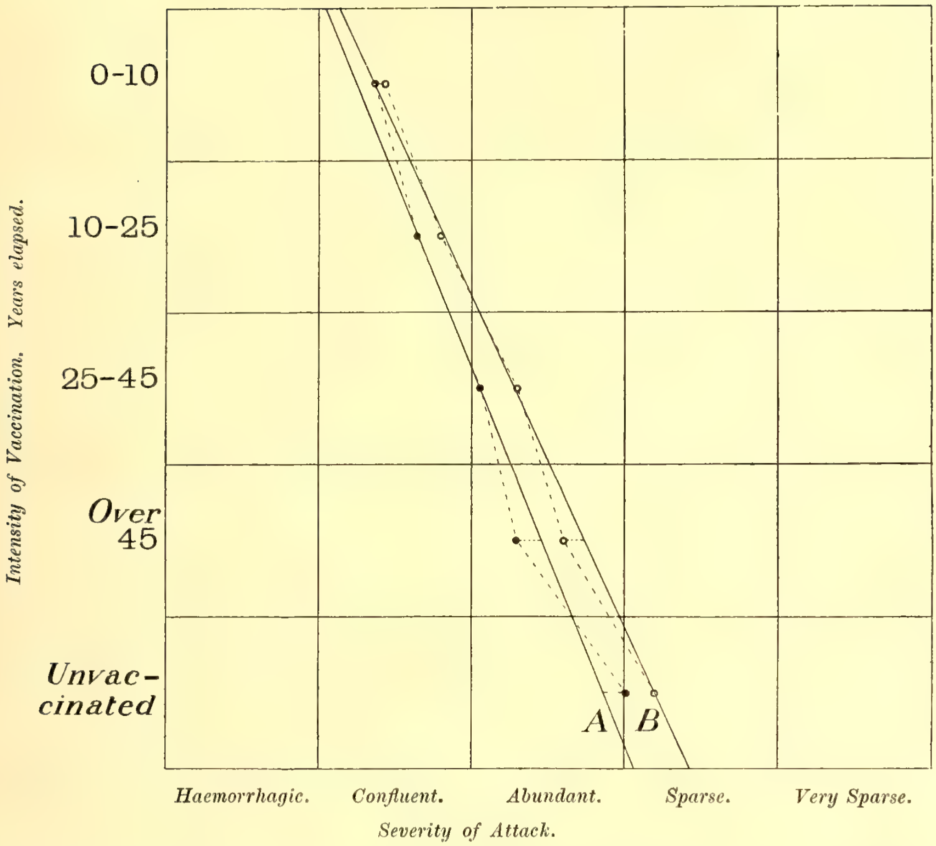


DIAGRAM XIII.

(ii)



changes would have been made had we used as horizontal scale a Gaussian calculated on the new distribution of total severities. It will be seen on comparing Diagrams XIV and XV that no substantial difference is made in the fundamental result. Further to test the effect of concentration we have had η the correlation ratio calculated by concentration at the means of the unselected scale of years since vaccination and by treating each array as a Gaussian; the value found for it in the former case (when uncorrected for class-index correlation of severity) is ·32, while in the latter case the three-rowed table gave $\eta = \cdot 32$ and ·34, with a mean value ·33*. Both these values would need the same class-index correction for severity, i.e. $r_{xC_x} = \cdot 9502$, which renders them ·34 and ·36 respectively.

Further, to show that a wide range of hypotheses can be made without modifying the fundamental conclusion, we have again plotted the regression lines on the Yulean hypothesis of pseudo-ranks before and after equalising deaths and

* *Biometrika*, Vol. VII. p. 257. The value of r found by concentrating at the Gaussian means of the marginal subranges is ·34, quite close to the η values as corrected.

recoveries, vaccinations and non-vaccinations (see Curves B of Diagrams XII and XIII). The fundamental result is obvious in all these diagrams, the selection, which does not change Mr Yule's coefficients, changes widely the true relationship as measured by the regression lines between the variates immunity conferred by vaccination and the severity of small-pox. *Take a given grade of severity of the disease, and the grade of immunity which is associated with it is entirely altered by the selection, or take a definite grade of vaccination and the average severity of the disease is represented by quite different figures before and after selection.* Mr Yule's most important property of his coefficients—the property that they remain unaltered by this selection—is the very property which forces us to the conviction that they are wholly unsuited to use in such a case as that which he asserts is eminently fitted for their use.

If we consider the matter algebraically, we find:

	η		r *		Yulean ϕ	Contingency
	Uncorrected	Corrected	Uncorrected	Corrected		
Before selection	·32 to ·34	·34 to ·36	·26	·34	·31	·36
After selection	·58	·61	·52	·60	·50	·58

We contend that a method which so substantially changes the real relationship between immunity and severity is wholly incapable of leading those who use it to any just inference as to the association of immunity and severity. It is only a logical quibble quite unworthy of the reputation of the man who uses it, and directed at an audience which had made no thorough study of the mathematics of statistics, to state as Mr Yule does that a man dies or does not die after incurring small-pox, and that he is vaccinated or not vaccinated. Death and vaccination are crude class-indices of severity and immunity, and Mr Yule's coefficients tell us nothing of the real relation between the two variates which is what we at least are seeking.

(13) *On the Stability of Coefficients of Association.*

If we have, as we believe we certainly have, in this paper succeeded in demonstrating the idleness of Mr Yule's coefficients of association and colligation for any purpose of applied statistics; if, as we hold, they are merely of interest from the standpoint of symbolic logic, i.e. in the discussion of *verbal* classifications, not in the treatment of the real things lying behind class-indices; the reader may naturally ask how we propose to treat similar problems. The answer to our mind is fairly clear. When the variates truly advance by unit grades, then there is no difficulty whatever about the problem. The right method to use is the product-

* Calculated by concentrating the contents of each cell at the Gaussian centroids of the subranges of the marginal totals and correcting for class-index correlations; see *Biometrika*, Vol. IX. p. 119.

moment correlation. This method has been used from the "earliest times" by biometricians, e.g. in dealing with the teeth on the carapace of prawns, the prickles on holly leaves, stigmatic bands on poppies, veins on leaves, etc., etc. There is absolutely nothing new in the Yulean method of pseudo-ranks applied to these data, the rank is the true measure of the variate in such cases. But the method is wholly false when applied to continuous variates in such cases as Mr Yule applies it. It is wholly false in particular when applied to fourfold tables, unless the difference between the two classes is a measurable unit. This, as we have indicated, is not the case in the Mendelian results of actual practice; it is not the case in the vaccination data or in any one of the cases to which Mr Yule has applied his coefficients. In no cases is he dealing with discrete unit differences. The nearest approach to a discrete difference is possibly in the case

DIAGRAMS XIV AND XV. Regression of Intensity of Vaccination upon Severity of Small-pox. (*A*) Original data. (*B*) Vaccinated made 50 °/₀ and haemorrhagic and confluent 50 °/₀ (Yule's Hypothesis). Intensity of Vaccination and Severity of Attack treated by Gaussian Methods.

(i) Assuming "severity" has a Gaussian distribution after the change of frequencies.

(ii) Assuming the centroids of each severity group unaltered by the change of frequencies.

DIAGRAM XIV.

(i)

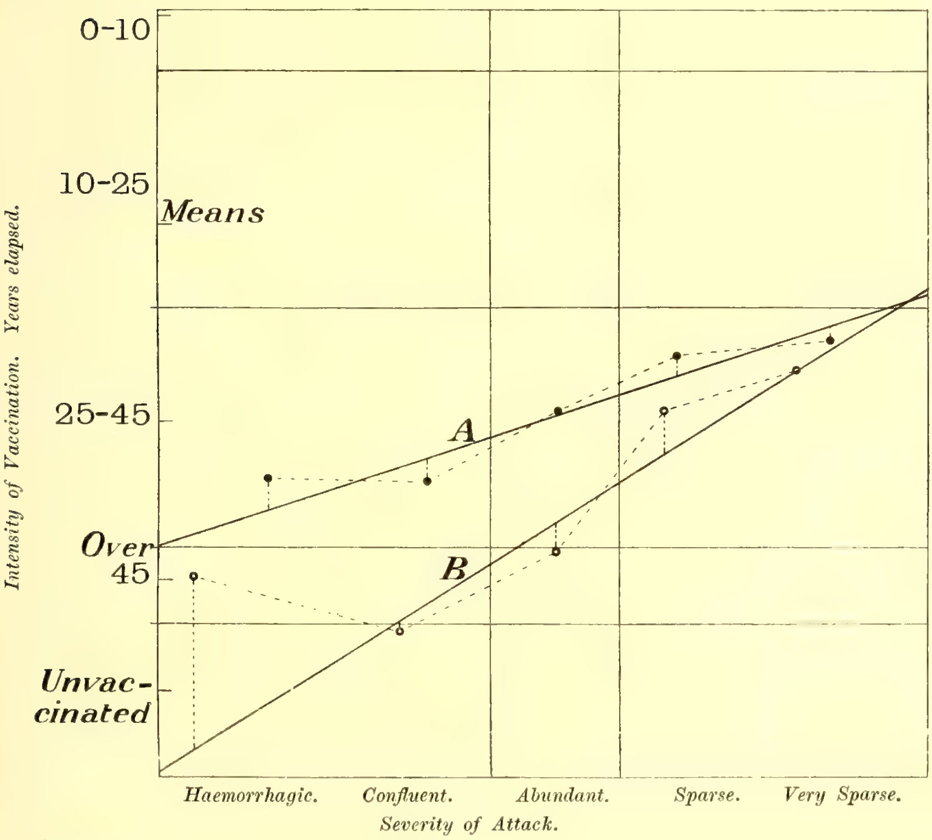
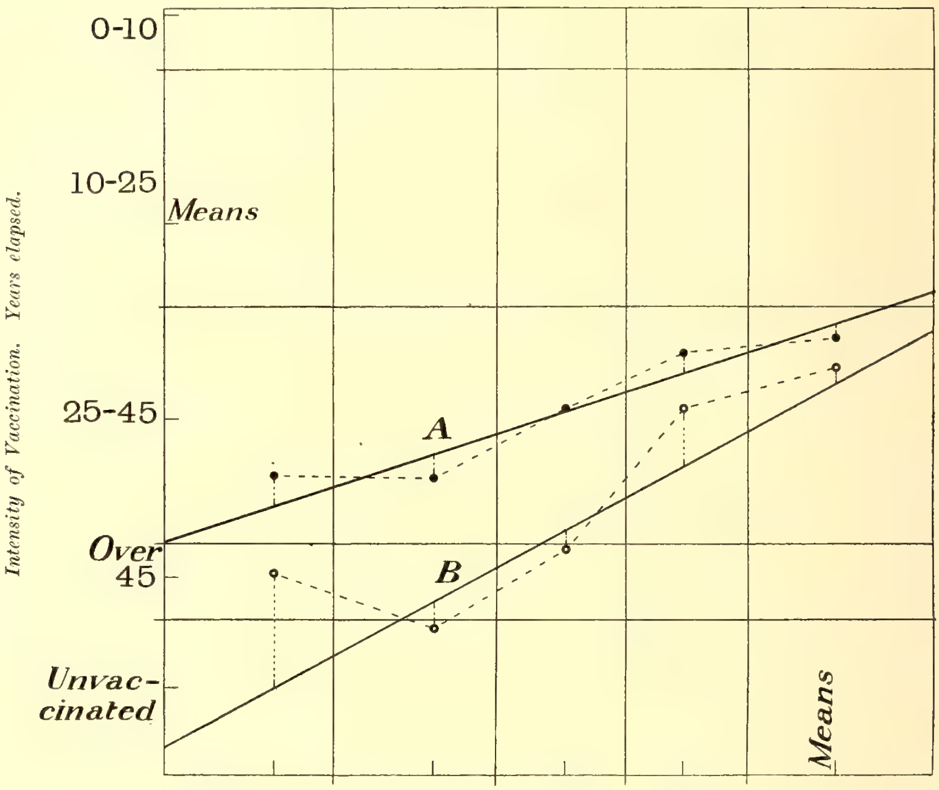


DIAGRAM XV.

(ii)



of sex, or in the use of a spear or a sword in negro tribes as alternatives, although even in these cases it is hard to understand why the male is a unit in excess or defect of the female or why a spear is a sword plus or minus a unit. But if we admit this, then ϕ is the appropriate coefficient to use, and not Mr Yule's association or colligation, and it should be applied to the original data and not to the adjusted or equalised table. Personally we should never use ϕ in such cases; we should measure the probability that the variates were independent, i.e. deduce the P from $n\phi^2$ by Elderton's tables, and this would guide our judgment in the matter*.

* We can get rid of the main effect of influence of total number of cases considered, and of the inequality of the marginal grouping, by adopting the method proposed by Pearson and thinking on a correlation scale: see "On a Novel Method of regarding the Association of two Variates classed solely in Alternate Categories," *Drapers' Company Research Memoirs*, Biometric Series VIII, Dulau & Co. 1912. In that paper the values of the correlation on the probability scale r_P were only tested against tetrachoric r_t for such totals and divisions as occur most frequently in everyday practice. A caution must be given here against the extension of that method without fuller investigation of such cases to divisions giving small percentages in the marginal frequencies. For such cases we know that tetrachoric r_t gives poor results.

But the cases of approach to discrete differences are exceedingly few in number compared to the mass of cases with which we have to deal; they exist as a rule only in the class-names, not in the things classified. Personally we have rarely found them except as already stated in *theoretical* Mendelian investigations, and to such cases the method of ranks, i.e. use of a ranks coefficient whether for fourfold or manifold tables, was first and rightly applied by Pearson. The Biometric School criticise not the application to Mendelian theory of these methods by Dr Brownlee and others, in which they only followed what had already been done, but Dr Brownlee's applying tetrachoric r_t to such theoretical tables and then supposing that he had got at the root of the difference between theoretical and actual heredity correlation tables! Mr Yule nowhere distinguishes between a theoretical Mendelian table and what we can absolutely demonstrate not to be theoretical Mendelism, e.g. Pearson's pigmentation tables. He writes: "As regards Dr Snow's recent comments in *Biometrika* on the use of the normal coefficient for Mendelian tables in Dr Brownlee's paper, he really thought that those comments were a much stronger condemnation of Professor Pearson's than of Dr Brownlee's work. Professor Pearson had repeatedly used the normal coefficient for inheritance tables, that were in all probability representations of Mendelian inheritance, as if it were an approximation to the product-sum correlation" (*loc. cit.* p. 651). It is a pity Mr Yule has not the courage of his opinions, and did not assert that the eye-colour data had a correlation of the Mendelian value $\frac{1}{3}$ because they *were* Mendelian. Instead of that he throws the sop of $\frac{1}{3}$ to the Mendelian Cerberus, having carefully produced it not from the proper fourfold Mendelian table, but by applying the method of pseudo-ranks, which involves just the same assumption of continuous variation and regression beyond the Mendelian divisions into unit characters, as is involved in the tetrachoric r_t. What Mr Yule means by "in all probability representations of Mendelian inheritance" is eloquently demonstrated by our discussion of the case—eye-colour in father and son—which he has himself selected to illustrate the Mendelian $\frac{1}{3}$.

If we leave the Mendelian *theoretical* table and any other *truly discrete* fourfold classifications, which are indeed difficult to find, and pass to the fourfold classification in general, what method are we to use? We assert that the fourfold or tetrachoric r_t is a coefficient of association infinitely superior to Mr Yule's old Q or new ω.

Mr Yule says he attaches no importance to the fluctuations of Q and ω for different divisions of the same table. If so, how can he usefully compare the values of Q or ω found from two tables with different divisions? If so, on what ground can he complain of the use of tetrachoric r_t because in certain cases it fluctuates for different divisions of the same table? He admits that "the coefficients of association and of colligation for different divisions of the same table in many cases fluctuate more largely than the normal coefficient" (*loc. cit.* p. 651), yet the only illustration he gives in his paper is that of the

table for eye-colour in brothers, for which he discusses the relative fluctuations of his coefficients and tetrachoric r_t by a method which from the standpoint of scientific statistics is wholly inadequate. He states that "as soon as we leave the narrow field within which normal or 'strained normal' correlation holds good, the normal coefficient fluctuates as we change the axes of division quite as largely as any other coefficient" (*loc. cit.* p. 633). How does this "narrow field" tally with the "many cases" in which these other coefficients fluctuate more largely? Mr Yule has purposely selected three or four markedly skew distributions to show how the tetrachoric r_t fluctuates; why did he not test adequately and completely its fluctuation as against those of the coefficient of association on this material? Would he ever have written his above dogmatic assertion had he done so?

Now we do not for a moment agree with Mr Yule that there is no importance in the fluctuations of a coefficient of association for different divisions of the same surface. On the contrary we assert that a true coefficient of association should be as stable as possible, that is to say that for a given surface it should have fluctuations which are within a reasonable range indicated say by twice its probable error. Indeed, for most practical purposes fluctuations of ·05 are of small importance. It is purely idle to do as Mr Yule has done, namely proceed to test such fluctuations by their ranges or their standard deviations obtained from the raw values. Each value must be considered in conjunction with its probable error. $Q = 1$ and $r = {\cdot}70$ are not comparable if the probable error of the first is zero and of the second ·05, the weight of the first observation is infinite and of the second finite. Mr Yule has compared such results without any regard to their relative weights.

Starting with tetrachoric r_t we have a definite surface, the Gaussian, either in its original or strained condition, for which there are no fluctuations in r_t, except such as might arise from random sampling. No such surface has been discussed by Mr Yule for Q or ϕ. But is an approximation to the Gaussian surface a rarity? Does it only occur in "a narrow field"? On the contrary it covers within the approximations required in practice a very wide field, namely nearly all distributions in anthropometry and many characters in plant, insect and animal forms. Even the irregularities of the eye-colour data may well turn out to be due to neglect of the age-corrections and not to failure of the Gaussian system. Can Mr Yule produce, even with careful seeking, actual distributions, in number one-tenth of the Gaussian cases just referred to, for which Q and therefore ω are practically constant for all divisions? There is no reasonable and logically consistent theory of deviations which leads to a surface of constant Q, although there is such a theory leading to a surface of constant r_t. We hold that stability of an association coefficient is not only a desirable, but an essential part of any true theory of association, and the fact that one theory does give a relatively stable coefficient for a large section of material is immensely in its favour.

But there are two other points about tetrachoric r_t which also in our opinion weigh much on its side when comparison is made with Q:

(i) It is modified by every form of selection and thus corresponds to our experience of every true measure of relationship. If association is to be of profit, it should pass into correlation as our detailed knowledge of the material becomes greater. All Mr Yule's association achieves is the passage with increasing knowledge into manifold and contradictory diversity. Every selection he has made on the basis of lesser knowledge must become an increased source of contradictory values, as he reaches more detailed knowledge of his material. There is a systematic variation of Q; it rises continuously from median to extreme divisions in every distribution on which we have tested it. That is to say, there is a wide divergence of practical statistical data from the surface of constant Q, and this in a definite direction. Nor is this to be wondered at, for the surface of constant Q has marked heteroscedasticity, and markedly curved regression: see Appendix III.

(ii) A true coefficient of association should not necessarily become perfect, when one of the four quadrants of the fourfold table becomes zero*. Whatever the degree of dependence of two variates may be, the frequency surface in practice is limited in extent, therefore by taking small percentages of one or both variates in the marginal totals we can always in practice make one quadrant zero. We hold indeed that a fitting coefficient of association need not necessarily be perfect, even if two opposite quadrants have zero frequencies†.

In both of these respects tetrachoric r_t is superior to the coefficients of association or colligation. The one aspect in which it is inferior is the labour of calculation, but the tables published by Everitt and an additional table which we hope shortly will be published render the labour by no means severe.

In order to establish our position it remains to show how even in the exceptional cases selected by Mr Yule—which do not represent the run of ordinary practice—the tetrachoric r_t is much more stable than his Q.

* Consider the following Table for true $r = {\cdot}52$:

Father's Stature.

Son's Stature.	Under *71·5*	Over *71·5*	Totals
Under *65·5* ...	127·5	0	127·5
Over *65·5* ...	863·5	87	950·5
Totals ...	991	87	1078

What real knowledge do we gain by saying that association is perfect between "tall" sons and "tall" fathers, when it arises solely from the extreme position of the father-division not giving any content to a second quadrant on a limited surface of imperfect correlation?

† See p. 177 and later Appendix I.

(A) *Barometric Data.*

The following table gives the value of tetrachoric r_t and of the coefficient of association Q*.

Southampton and Laudale.

Divisions	Tetrachoric r_t	Yule's Association	Divisions	Tetrachoric r_t	Yule's Association
29·15	·6969 ± ·0378	·9371 ± ·0144	*29·95*	·7991 ± ·0096	·8868 ± ·0075
29·25	·7254 ± ·0291	·9335 ± ·0124	*30·05*	·7954 ± ·0100	·8831 ± ·0075
29·35	·7476 ± ·0220	·9244 ± ·0108	*30·15*	·7971 ± ·0110	·8867 ± ·0075
29·45	·7410 ± ·0215	·9067 ± ·0109	*30·25*	·8112 ± ·0127	·9109 ± ·0070
29·55	·7303 ± ·0177	·8847 ± ·0111	*30·35*	·7983 ± ·0149	·9229 ± ·0076
29·65	·7559 ± ·0140	·8866 ± ·0097	*30·45*	·8300 ± ·0176	·9579 ± ·0061
29·75	·7848 ± ·0114	·8951 ± ·0083	*30·55*	·8452 ± ·0231	·9761 ± ·0050
29·85	·7917 ± ·0102	·8874 ± ·0079	—	—	—

* We have not thought it needful to deal in every case with *both* Mr Yule's coefficients. Since ω is a function of Q, and its probable error is a function of the probable error of Q, there is no occasion to do so, and it opens a whole field for the display of statistical fallacy. Consider two quantities r and ρ defined in their mutual relations by $r=(1-\epsilon)/(1+\epsilon)$ and $\rho=(1-\epsilon^{1/n})/(1+\epsilon^{1/n})$; then if we arrange a fourfold table so that r is positive, ϵ will lie between 0 and 1 and ρ between 0 and 1. Clearly

$$\epsilon=(1-r)/(1+r), \qquad \epsilon^{1/n}=(1-\rho)/(1+\rho),$$

or taking logarithmic differentials we find

$$\sigma_\epsilon/\epsilon=2\sigma_r/(1-r^2) \quad \text{and} \quad \frac{1}{n}\sigma_\epsilon/\epsilon=2\sigma_\rho/(1-\rho^2).$$

Thus it follows that
$$\frac{\sigma_\rho}{\rho}=\frac{\sigma_r}{r}\times\left(\frac{1}{n}\frac{\frac{1}{\rho}-\rho}{\frac{1}{r}-r}\right).$$

Now if n be positive r is always greater than ρ; it follows accordingly that $1/\rho-\rho$ is greater than $1/r-r$, but if we have any given series we can choose the value of n so great that σ_ρ/ρ is less than σ_r/r. For if $u=\dfrac{\sigma_\rho/\rho}{\sigma_r/r}$, then $u=\dfrac{1}{n}\dfrac{\epsilon^{1/n}}{1-\epsilon^{2/n}}\dfrac{1-\epsilon^2}{\epsilon}$. As n becomes large, then the limit of

$$\frac{1}{n}\frac{\epsilon^{1/n}}{1-\epsilon^{2/n}}=\frac{1}{2n},$$

and therefore
$$\text{Limit of } u=\frac{1-\epsilon^2}{2n\epsilon},$$

which can be made as small as we please.

Thus if we have any system of tetrachoric r_t's, it is always possible to devise a new system of ρ's, obtainable by the relations

$$\epsilon=(1-r)/(1+r) \quad \text{and} \quad \rho=(1-\epsilon^{1/n})/(1+\epsilon^{1/n}),$$

which have σ_ρ/ρ less than σ_r/r. Thus, if we generalise Mr Yule, $\Omega=(1-\kappa^{1/n})/(1+\kappa^{1/n})$ satisfies all his conditions of a coefficient of association, but we can always select n so that σ_Ω/Ω shall be less than σ_Q/Q or σ_ω/ω, where Q and ω are Mr Yule's coefficients. Similarly we can deduce from the tetrachoric r_t a definite function of it ρ, which has a ratio σ_ρ/ρ less than σ_r/r, and yet satisfies all the so-called conditions of association. But for Mr Yule, and many others, the ratio of a quantity to its probable error is a measure of its significance. Hence by a proper choice of n we can always

It will be seen that the values we have obtained are not wholly in agreement even to the second place of figures with the values obtained by Mr Yule, who has selected only 8 out of the 15 diagonal cases. Diagrams XVI and XVII show that

DIAGRAM XVI. Correlation of Barometer at Laudale and Southampton. Actual values of tetrachoric r_t.

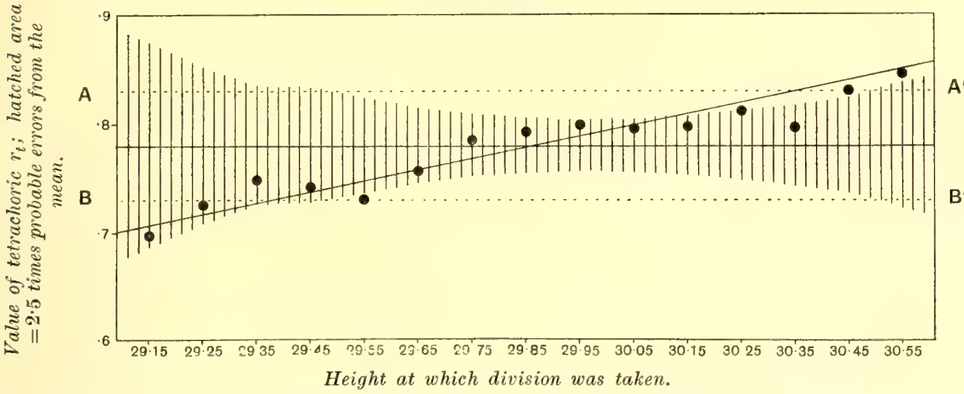
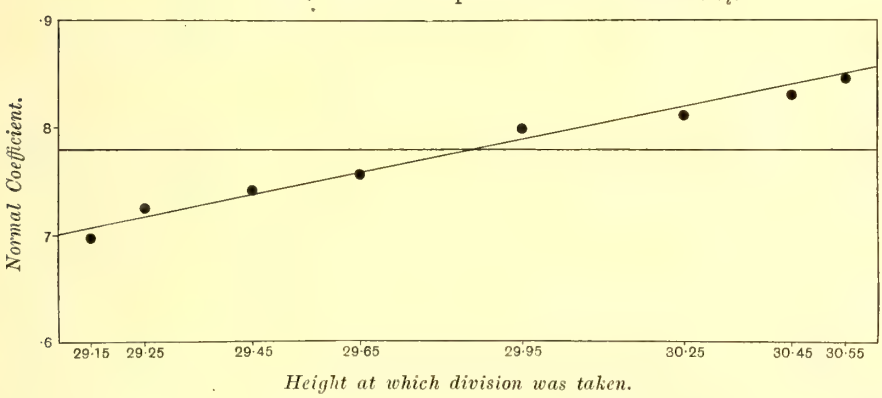


DIAGRAM XVII. Mr Yule's representation of tetrachoric r_t.



deduce from one of Mr Yule's type of coefficients a second which has greater or less or *no significance* as the case may be. Further:

$$\sigma_\rho = \frac{1}{n}\,\frac{1-\rho^2}{1-r^2}\,\sigma_r = u\,\sigma_r,$$

where $u = \dfrac{1}{n}\,\dfrac{\epsilon^{1/n}}{(1+\epsilon^{1/n})^2}\,\dfrac{(1+\epsilon)^2}{\epsilon}$, but as n increases this tends to take the limit:

$$\frac{1}{n}\,\frac{(1+\kappa)^2}{4\kappa},$$

or can be made less than unity. In other words, given r or Q, we can always choose a new quantity ρ or Ω, which has a smaller absolute probable error. If colligation for a given series has a smaller probable error than Q or a smaller one than r_t, it is always possible to choose new functions ρ or Ω, which are absolutely determined by r_t or by Q and ω, which have still smaller probable errors, and yet are true Yulean coefficients of association, ranging between 0 and 1! The fact is that the term "probable error" has no meaning for these coefficients until Mr Yule has discussed the nature of their frequency distribution, which is certainly not Gaussian. In the text we have compared r_t and Q with their probable errors because it probably gives some rough estimate of their relative stability, and it is the test Mr Yule has himself chosen.

Mr Yule's selection has been singularly favourable to his unjustified assumption that the points lie along a straight line; actually the points are extremely steady between 29·85 and 30·35, and tail off to the terminals, where the percentages in the smallest quadrant are much smaller. However, apart from the question of linearity of the distribution of r_t there is little doubt that the values are in defect for low divisions and in excess for high divisions. The question therefore is, are these defects and excesses such as to invalidate the use of the method of tetrachoric functions applied even to variates as skew as the barometer data? The only method of answering this question appears to us to consider the relation of the values found to their probable errors, and again the amount of stress which is likely in practice to be laid on the result deduced from a single fourfold table. We believe that practically it is not necessary for the type of reasoning based on a fourfold table that the correlation found should be nearer than $\pm$ ·05. In the present case three coefficients exceed these limits, but if we proceed solely to two figures the first and last only lie outside these limits. Diagram XVI shows this result by the lines AA' and BB'. The hatched part of the diagram corresponds to the region on either side the product-moment value of 2·5 times the probable error. Actually some 10°/ₒ of the observations should lie outside these limits, i.e. $1\frac{1}{2}$ observations instead of 3. But even these three have almost contact with the hatched area. It seems to us that in this very case of barometer data, which Mr Yule has chosen for its marked skewness to discredit tetrachoric r_t, the coefficient defeats Mr Yule completely!

Now let us compare the results with Mr Yule's own coefficient of association. The difficulty in the comparison lies with the standard value of Q against which the other Q's are to be compared. Had Mr Yule studied the surface of constant Q—what we may term the association-surface—(see p. 184 above), then the Q corresponding to the best fitting association surface would have formed a standard Q. But we have at present no means of finding this standard Q, and Mr Yule tells us that he himself lays no stress on the diversity of values obtained for Q with different divisions. However, Mr Yule has himself taken the mean Q as a standard when he comes in a special case to deal with the relative stability of tetrachoric r_t and association Q. Accordingly we follow him in taking a mean Q. But he has gone astray in simply taking the arithmetical average of his Q's or r_t's without regard to their probable errors. The proper means to take are *weighted* means, and the proper standard deviations are weighted standard deviations. Each value must be weighted with the inverse square of its probable error as the measure of its grade of accuracy. With this weighting we find:

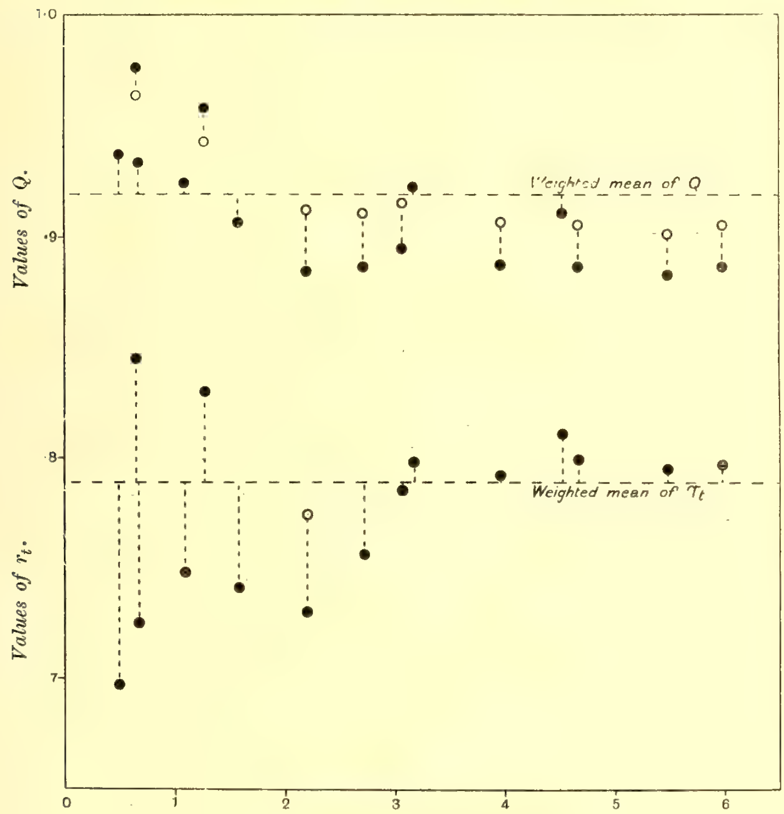
Barometric Heights.

	Weighted Mean	Weighted Standard Deviation
Tetrachoric r_t ...	·7886	·0243
Association Q ...	·9194	·0346

It will be seen that the mean weighted tetrachoric r_t differs only by ·0086 from the true product-moment value (·7800). Further the variability of the association coefficient marks an increase of 42·4 °/。 on the variability of the tetrachoric r_t*; there can be no doubt that in this first of Mr Yule's selected markedly skew cases the tetrachoric coefficient is far more stable than the association coefficient. Another method of approaching the problem will also illustrate this point. Let us express the deviations from the means in the case of both coefficients in terms

DIAGRAM XVIII. Comparison of r_t and Q for Correlation of Barometric Heights. Q has nine cases, r_t only one case outside 2·5 times the probable error.

Barometer Heights. Southampton and Laudale.



Percentage in smallest quadrant.

of their probable errors, and then average these deviations: the mean value of $(Q - \text{mean } Q)/(\text{probable error of } Q) = 3·35$, the mean value of $(r_t - \text{mean } r_t)/(\text{probable error of } r_t) = 1·63$. Clearly Q is 100 °/。 worse than r_t. This is indicated on the accompanying Diagram XVIII, where the values of Q and r_t are plotted to the percentage frequency of the smallest quadrant of the table, a matter which

* As both Q and r_t have the same possible range of variation -1 to $+1$, in such a case the standard-deviation and *not* the coefficient of variation properly measures the variability.

we have found a safe guide to determine how far we introduce risk in using tetrachoric r_t. It will be seen that the unsatisfactory values of r_t occur when this smallest quadrantal frequency is less than 2·5 °/$_{\circ}$. On the side of Q or r_t, towards their mean, a line is drawn equal to 2·5 times the probable error; in all but one case of r_t this reaches the mean; in nine cases of Q it does not, and the deviations in six of these cases are excessive. There can, we think, be no doubt that tetrachoric r_t is in this first of Mr Yule's selected cases far more stable than his coefficient of association.

(B) *Skew Table of Number of Mendelian Couplets in Father and Son* (see Table XVI on our p. 222).

The following table gives tetrachoric r_t and Q with their probable errors for a second series selected by Mr Yule.

Vertical division taken between

	0—1	*1—2*	*2—3*	*3—4*
0—1	r_t = ·39 ± ·017 Q = ·49 ± ·018	r_t = ·37 ± ·017 Q = ·51 ± ·023	r_t = ·35 ± ·030 Q = ·63 ± ·048	r_t = ·33 ± ·086 Q = ·75 ± ·153
1—2	r_t = ·37 ± ·017 Q = ·51 ± ·023	r_t = ·40 ± ·017 Q = ·52 ± ·019	r_t = ·39 ± ·025 Q = ·61 ± ·029	r_t = ·37 ± ·066 Q = ·72 ± ·087
2—3	r_t = ·35 ± ·030 Q = ·63 ± ·048	r_t = ·39 ± ·025 Q = ·61 ± ·029	r_t = ·41 ± ·033 Q = ·69 ± ·034	r_t = ·40 ± ·073 Q = ·79 ± ·068
3—4	r_t = ·33 ± ·086 Q = ·75 ± ·153	r_t = ·37 ± ·066 Q = ·72 ± ·087	r_t = ·40 ± ·073 Q = ·79 ± ·068	r_t = ·43 ± ·122 Q = ·90 ± ·072

It is interesting to note the high values reached by the probable errors of tetrachoric r_t for the small quadrant frequency divisions. Both systems of probable* errors are based on the assumption that the total frequency is 4096.

The following results were obtained:

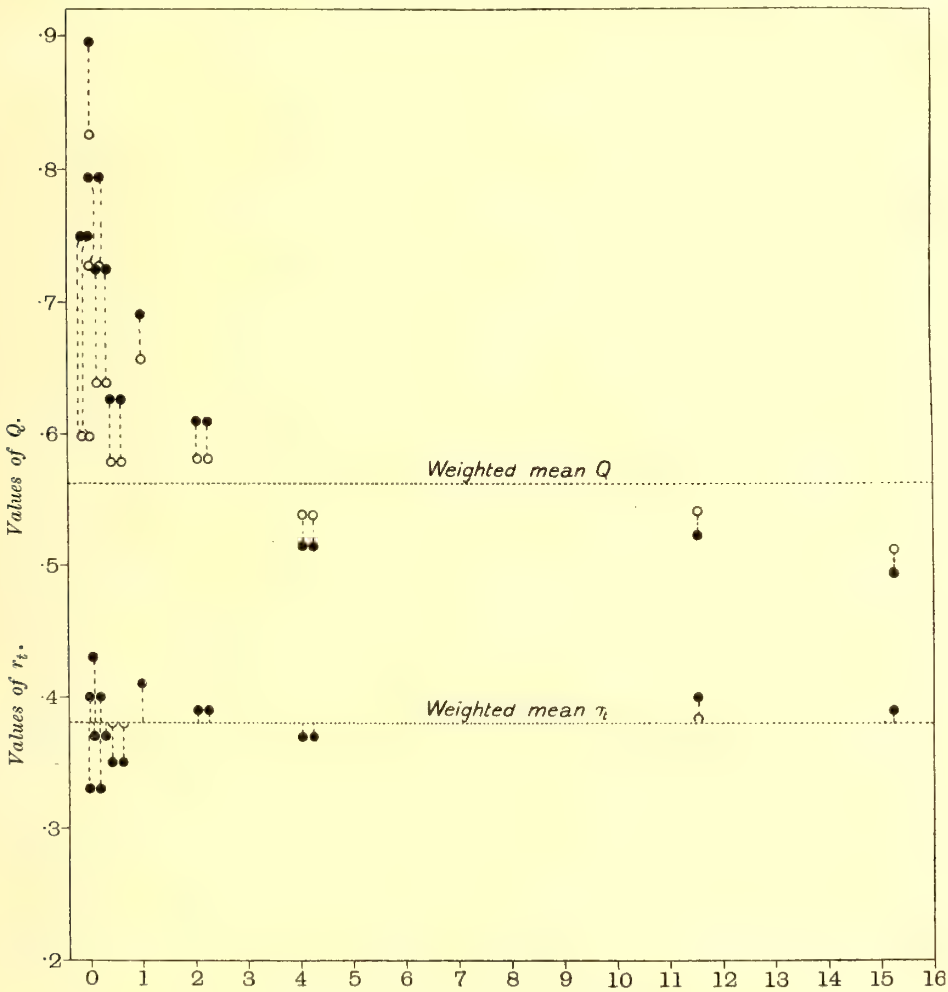
	Weighted Mean	Weighted Standard Deviation
Tetrachoric r_t ...	·3809	·0176
Association Q ...	·5627	·0840

Of course this table of discrete couplets is not one to which we should think of applying tetrachoric r_t, but Mr Yule has applied it; however, the weighted

* In using probable errors at all here for a basis of comparison, we assume that the table may be considered to have arisen from actual observations.

mean of such method when applied is only ·048 in error, the contingency method giving ·3288 and the true value being ·3333. We see that, judged by weighted standard deviations, the relative variabilities are as 176 to 840, or the stability of the tetrachoric r_t is 4·8 times as great as that of Mr Yule's coefficient.

DIAGRAMS XIX AND XX. No Q differs by less than its probable error from the weighted mean, only one r_t differs by more than its probable error.



We can look at this from the standpoint of the diagrammatic representation of the probable error (see Diagrams XIX and XX). In the lower figure we have the 16 points given by tetrachoric r_t; *only a single* one (shown by the individual linked with the non-black circle) is at a distance more than *once* the probable error from the weighted mean. In the upper figure *not a single case* occurs in which Q differs by so little as its probable error from the weighted mean value.

34—2

There is indeed no comparison between the stability of tetrachoric r_t and Q in an illustration which Mr Yule himself has selected to indicate the badness of tetrachoric r_t! Mr Yule's coefficient of association, we are told by the statisticians of the Royal Statistical Society, will be used for Mendelian problems as soon as the Mendelians know a little algebra, but here in the very case of Mendelian data we find, as a mere coefficient of association, the r_t of the despised Gaussian theory is immensely superior to the Yulean coefficient*. Judged algebraically:

$$\text{Average value of } (Q - \text{mean } Q)/(\text{probable error of } Q) = 2{\cdot}34,$$

$$\text{Average value of } (r_t - \text{mean } r_t)/(\text{probable error of } r_t) = {\cdot}57,$$

or Q is four times as unstable as tetrachoric r_t.

(C) *Severity of Small-Pox and Strength of Vaccination Immunity.*

Sixteen cases can be worked out for tetrachoric r_t and Q for this material, and there is a great variation in both r_t and Q†.

	Haemorrhagic and Confluent	Confluent and Abundant	Abundant and Sparse	Sparse and Very Sparse
At *10*	$r_t = ? \pm ?$ $Q = 1 \pm {\cdot}0000$	$r_t = {\cdot}3602 \pm {\cdot}0911$ $Q = {\cdot}7892 \pm {\cdot}1296$	$r_t = {\cdot}2857 \pm {\cdot}0614$ $Q = {\cdot}5415 \pm {\cdot}1036$	$r_t = {\cdot}2226 \pm {\cdot}0626$ $Q = {\cdot}4319 \pm {\cdot}1036$
At *25*	$r_t = {\cdot}2905 \pm {\cdot}0598$ $Q = {\cdot}5711 \pm {\cdot}1079$	$r_t = {\cdot}3694 \pm {\cdot}0281$ $Q = {\cdot}5411 \pm {\cdot}0413$	$r_t = {\cdot}3484 \pm {\cdot}0249$ $Q = {\cdot}4469 \pm {\cdot}0303$	$r_t = {\cdot}2500 \pm {\cdot}0293$ $Q = {\cdot}3444 \pm {\cdot}0369$
At *45*	$r_t = {\cdot}2187 \pm {\cdot}0556$ $Q = {\cdot}4111 \pm {\cdot}0890$	$r_t = {\cdot}4169 \pm {\cdot}0309$ $Q = {\cdot}5714 \pm {\cdot}0326$	$r_t = {\cdot}3961 \pm {\cdot}0275$ $Q = {\cdot}5474 \pm {\cdot}0351$	$r_t = {\cdot}2578 \pm {\cdot}0332$ $Q = {\cdot}4143 \pm {\cdot}0569$
Between *45* and Unvaccinated	$r_t = {\cdot}0426 \pm {\cdot}0770$ $Q = {\cdot}1009 \pm {\cdot}1772$	$r_t = {\cdot}5121 \pm {\cdot}0363$ $Q = {\cdot}7220 \pm {\cdot}0322$	$r_t = {\cdot}6022 \pm {\cdot}0287$ $Q = {\cdot}8599 \pm {\cdot}0308$	$r_t = {\cdot}5381 \pm {\cdot}0349$ $Q = {\cdot}8862 \pm {\cdot}0524$

In the form in which Table XXII is given on our p. 254, all values of Q and r_t are *negative.*

This table gives us the following results:

	Weighted Mean	Weighted Standard Deviation
Tetrachoric r_t ...	·3827	·1211
Association Q ...	·5902	·1737

* It is needless to say we should never have thought of applying either of these coefficients to theoretical Mendelian data; we hold that the correct method was the method applied to this very case by one of us *ab initio*, namely that of product-moment r.

† There can be little doubt that the extreme variations are due to 9 out of the 16 divisions giving a quadrant of minimum frequency with less than 1 °/₀ of frequency in it.

In forming these weighted means we have left out the value of $Q = 1 \pm \cdot 0000$ and of $r_t = ? \pm ?$*. In the first place had we included this value of Q, its weight would have been infinite, and whatever the value of r_t and its probable error may be, it follows that Q must be more unstable than r_t. We have got rid of this anomalous Q, although Mr Yule's theory gives it infinite weight, and dealt only with the remaining 15 cases; but to do this is like considering the assets of a bankrupt, after we have disregarded the claims of his principal creditor. The above table, however, shows that Q is, notwithstanding this disregard, much worse than r_t; in fact 43·4 °/ₒ must be added to the variability of r_t to reach that of Q.

If we measure the deviations from the weighted means in terms of the probable errors as before we find for tetrachoric r_t: the mean $= 2{\cdot}71$ and for association Q: $= 3{\cdot}31$. Thus again, although since $Q = 1$ is excluded, we find to a lesser extent tetrachoric r_t more stable than Q.

(D) *Lengths of Ivy Leaves.*

Mr Yule directly selected the lengths of growing ivy leaves as an especially skew distribution upon which to test tetrachoric r_t, and this in a case where the table had been deduced for homotyposis, i.e. with all the local lumpiness which arises from that method of treatment. If we take the 4×4-fold of our Table XVII on p. 222, nine values of Q and tetrachoric r_t are available. They are:

	1—2	*2—3*	*3—4*
1—2	$r_t = \cdot 6998 \pm \cdot 0046$ $Q = \cdot 8531 \pm \cdot 0027$	$r_t = \cdot 6406 \pm \cdot 0039$ $Q = \cdot 8655 \pm \cdot 0040$	$r_t = \cdot 5572 \pm \cdot 0058$ $Q = \cdot 9167 \pm \cdot 0085$
2—3	$r_t = \cdot 6406 \pm \cdot 0039$ $Q = \cdot 8655 \pm \cdot 0040$	$r_t = \cdot 5731 \pm \cdot 0033$ $Q = \cdot 6768 \pm \cdot 0033$	$r_t = \cdot 5218 \pm \cdot 0046$ $Q = \cdot 7570 \pm \cdot 0048$
3—4	$r_t = \cdot 5572 \pm \cdot 0058$ $Q = \cdot 9167 \pm \cdot 0085$	$r_t = \cdot 5218 \pm \cdot 0046$ $Q = \cdot 7570 \pm \cdot 0048$	$r_t = \cdot 5548 \pm \cdot 0052$ $Q = \cdot 7508 \pm \cdot 0045$

We have the following results:

	Weighted Mean	Weighted Standard Deviation
Tetrachoric r_t ...	·5920	·0570
Association Q ...	·8024	·0757

* On the difficulty, in fact idleness, of attempting to determine r_t from tables with zero frequency in one quadrant: see Appendix I. That quadrant as we there show might have ·5 in it, or, indeed, the material from which it is drawn unity. In the former case 20 terms in the r_t-equation will not suffice to determine the value of r_t. In the latter case r_t might swing over from -1 to a small positive quantity. In fact the process of finding tetrachoric r_t in such cases warns its user that it is indeterminable. Users of Q will assert that the relationship is perfect with a zero probable error!

Therefore Q exhibits an increase of about 33 °/₀ on the variability of tetrachoric r_t. If we take the probable error test

$$\text{Average value of } \frac{r_t - \text{weighted mean } r_t}{\text{probable error}} = 11{\cdot}53,$$

$$\text{Average value of } \frac{Q - \text{weighted mean } Q}{\text{probable error}} = 16{\cdot}19.$$

Thus Q represents an increase of 40 °/₀ on the variability of r_t.

(E) *Correlation of Hair and Eye Colours. Livi's Data*†.

The table is:

Eye-Colour.

Hair-Colour.	Blue	Grey	Brown	Black	Totals
Blond ...	9083	8187	7031	217	24518
Red ...	343	518	819	37	1717
Brown ...	17829	39467	117522	4945	179763
Black ...	3627	13433	54883	20919	92862
Totals ...	30882	61605	180255	26118	298860

This table gives the following nine comparative results:

	Blue-Grey	Grey-Brown	Brown-Black
Blond-Red ...	$r_t = {\cdot}5307 \pm {\cdot}0026$ $Q = {\cdot}7442 \pm {\cdot}0023$	$r_t = {\cdot}5074 \pm {\cdot}0023$* $Q = {\cdot}7129 \pm {\cdot}0023$	$r_t = {\cdot}4392 \pm {\cdot}0054$* $Q = {\cdot}8422 \pm {\cdot}0067$
Red-Brown ...	$r_t = {\cdot}5239 \pm {\cdot}0030$ $Q = {\cdot}7356 \pm {\cdot}0023$	$r_t = {\cdot}5255 \pm {\cdot}0023$* $Q = {\cdot}7263 \pm {\cdot}0023$	$r_t = {\cdot}4332 \pm {\cdot}0048$* $Q = {\cdot}8294 \pm {\cdot}0067$
Brown-Black ...	$r_t = {\cdot}3601 \pm {\cdot}0025$ $Q = {\cdot}5791 \pm {\cdot}0041$	$r_t = {\cdot}3240 \pm {\cdot}0020$ $Q = {\cdot}4393 \pm {\cdot}0026$	$r_t = {\cdot}6449 \pm {\cdot}0020$ $Q = {\cdot}8365 \pm {\cdot}0016$

Now it is of interest to compare the values of r_t found by fourfold tables with those obtained by other methods. Mr Yule has nothing but the method of pseudo-ranks to apply to the table in its detailed form. This gives a Yulean ·3680; the value found by corrected mean square contingency is ·5189; the weighted mean of the nine tetrachoric r_t values is ·4842, much nearer to the contingency value. Indeed the Yulean, if corrected for ranks becomes ·3953, and if corrected for class-index correlations becomes ·5051, i.e. differs quite insignificantly

† *Antropometria militare*, Part I. p. 62.

* Mr Yule's values do not agree to the second decimal place with ours in these cases.

from the contingency ·5189*. It is thus easy to see why Mr Yule gets such an absurdly low value. If it be said that two out of the nine values of tetrachoric r_t are lower than the Yulean, the reply is a simple one: They are not divisions which would be made with our present knowledge of hair and eye pigmentation. There is no such thing as a "black" eye, at best it is only a dark shade of brown, and the division between brown and black is largely a matter of personal equation. Again the relative amount of pigmentation in a true grey eye may be as small as in many so-called blue eyes. The natural physiological division is between the blue-grey and the brown-black groups, although, if Livi put his hazels and greys with some pigment partly into blues and partly into browns, a division between blue and grey might give as good results. Passing now to the hair-colour, we believe the distinction between brown and black to be again a matter of personal equation, the shades of brown range up to black. Again red is a tint which may contain less pigment granules than the blond†, but in some varieties it has more granules than many browns. Hence the difficulty of such a division as that betwixt brown and red. The real division would come between blond and brown with reds omitted or better still microscopically tested and placed in their appropriate division. Anyhow the divisions corresponding to the four left-hand top quadrants of the last table seem to us the most satisfactory, and they all agree in giving the correlation of hair and eye colours = circa ·52, a result in excellent accord with the mean square contingency.

Pearson's data for British school children give by corrected contingency ·52, and the following table shows that the values for all nations must be in excess of that determined by Mr Yule's method:

Correlation of Hair and Eye Colours by corrected Contingency.

Italian Recruits	·519
British School Children	·524
Baden Recruits	·484
German Jewish Children	·444
Prussian School Children	·401
Swedish Recruits	·414
Italians by Yule's Method of Pseudo-Ranks ...	·368

* Assuming normal correlation, the correction for ranks is $r = 2\sin(30° \times ·3680) = ·3953$. For hair groups the class-index correlation is ·8686 and for eyes ·9010, and variate correlation

$$= ·3953/(·8686 \times ·9010) = ·5051.$$

Of course these are only rough approximations, for Mr Yule's theory of unit subranges precludes our using the corrections actually given. But any corrections will be roughly of this order.

† Some reds lack all granular pigmentation; there are blonds which also do this, but they are very few in number. Both reds and blonds without melanin pigment granules are truly albinotic hairs, whether accompanied or not by albinotic eyes.

Returning to the relative values of tetrachoric r_t and association Q we find:

	Weighted Mean	Weighted Standard Deviation
Tetrachoric r_t ...	·4894	·1102
Association Q ...	·7258	·1161

The variation is high for both coefficients, but even here tetrachoric r_t is better than Q, you must raise the variability of r_t 5·4 °/。 to reach the variability of Q. But the worth of the two coefficients is wholly unequal. If we are told the correlation of hair and eye colours is about ·50, a whole series of ideas is associated with this number; a very little additional information gives the relative standard deviations of the two variates—accurately enough for practice—and we picture to ourselves the regression lines and the associated changes in pigmentation of hair and eye classes. But what does an association coefficient of ·73 for special dichotomies tell us? We venture to assert that it conveys no information whatever to the investigator's mind, and is absolutely incomparable with other association coefficients of the same table, because each for the same system depends on the values of the variates at which the divisions are made, and because it has not the least relation with any physical properties of the distribution. If Mr Yule replies to these criticisms, that tetrachoric r_t is also unstable, if not to the same extent, and that a function of Q does provide (if the table be doctored) a certain difference of percentages, we answer that tetrachoric r_t is far from so unstable for the distributions of ordinary practice as he has endeavoured to make it out by selecting: (i) pigmentation data, which as long ago as 1901 were recognised as irregular, and (ii) markedly skew frequencies*. The instability is many times compensated by the definite physical significance of the coefficient.

Further the percentages which Mr Yule's deduced coefficient of colligation represents are wholly artificial and incapable of any rational interpretation. If we were to equalise the number of vaccinated and unvaccinated in any locality, how could we equalise the number of deaths and the number of recoveries, and what intelligible meaning can be given to the percentages when it has been done? How can we possibly give any interpretation to the result reached by a disciple of Mr Yule that the "index de corrélation"—i.e. Mr Yule's Q—between the stature of recruits and rent in the 20 districts of the city of Paris is "perfect"? Average stature of recruits for different districts forms a continuous variate system, so does average rent, and the two properly correlated would show the nature of the regression line, but this disciple of Mr Yule's, in order to save a little absolutely

* Mr Yule (*loc. cit.* p. 624) speaks of the selection he has made as "exhibiting moderately skew distributions." Unless he means that they are not U- or J-shaped curves, we consider this an entire misnomer.

needful arithmetic, tells us that the index of correlation between stature and rent is perfect*!

(F) *Eye-Colour Data of Pearson.*

We now come to the eye-colour data. These are the only cases we have so far worked out in which the variability of Q is at all comparable with that of tetrachoric r_t, and the reason is not very far to seek. We have already pointed out that the true difficulty in these eye-colour tables appears to lie in lumps of excess frequency in or near the corner cells of the quadrants of less frequency and that thus the material is heterogeneous in character. We think it quite probable that this is due to the inclusion of really senile fathers on the one hand and of infant sons on the other, or of pairs of brothers or sisters one of whom is an infant. In this manner eye tints which are originally, or will become, mediocre in colour may be classified as very light. The record which Sir Francis Galton provided gave no ages; the original data are now in the possession of the Galton Laboratory and it is proposed to reconsider both these and the Huxley Lecture data for eye-colour†, paying attention to change of eye-colour with age. This investigation will necessarily take a considerable time and it might be wiser to await its conclusion before entering further on this topic. But we should no doubt be told that we were omitting just those cases that appeared favourable to Mr Yule's association coefficient, and accordingly we have included the eye-colour data here. The table on p. 274 gives the Brother-Brother coefficients for 16 divisions.

Now this is a symmetrical table and Mr Yule reckons the diagonal coefficients once, and the repeated coefficients *twice*, but we are doubtful of the accuracy of this process. Such a symmetrical table leads to exactly the same results from symmetrically placed divisions, and it is not clear why double weight should be given to such a division as 2—3 and 6—7 because as the table is written out it occurs *twice*. It seems to us that the diagonal values and all on one side of them are the only independent coefficients; we have lost the independence of the coefficients in the cells symmetrically situated with regard to the diagonal by the very process of adding the tables for First Brother in terms of Second Brother

* A. Niceforo, "Contribution à l'étude des corrélations entre le bien-être économique et quelques faites de la vie démographique," *Journal de la Société de Statistique de Paris*, 52 Année (1911), pp. 322—341. Professor Niceforo studies "corrélations" by aid of the coefficient of association, which he applies to the following continuous variates: stature, rent, probable income, numbers of illiterates, of insanitary dwellings, of paupers, of pauper funerals, of workers, size of families, numbers of cubic feet of air space, of inhabitants to the acre, special death-rates from all sorts of diseases, general death-rates and birth-rates etc., etc. We believe that the whole of this work must be redone. Even if a rough estimate had been required, Mr Yule's coefficient should not have been used, but the division made at the median values and Sheppard's formula for tetrachoric r_t adopted. There is not even the excuse of apparent discreteness in any of Professor Niceforo's attributes.

† Pairs of siblings in the same school are much more nearly of an age than any pair between the ages of 5 and 15 taken from the population of school children. While the head-measurements were corrected for age, the hair and eye colours were not, and probably the changes are more important than we believed at that date.

First Brother.

Second Brother.	*2—3*	*3—4*	*4—5*	*6—7*
2—3	$r_t=\cdot5058\pm\cdot0172$ $Q=\cdot6214\pm\cdot0176$	$r_t=\cdot3761\pm\cdot0184$ $Q=\cdot4920\pm\cdot0230$	$r_t=\cdot2428\pm\cdot0212$ $Q=\cdot3434\pm\cdot0304$	$r_t=\cdot2373\pm\cdot0240$ $Q=\cdot3653\pm\cdot0383$
3—4	$r_t=\cdot3761\pm\cdot0184$ $Q=\cdot4920\pm\cdot0230$	$r_t=\cdot5185\pm\cdot0160$ $Q=\cdot6229\pm\cdot0165$	$r_t=\cdot3524\pm\cdot0196$ $Q=\cdot4603\pm\cdot0232$	$r_t=\cdot2777\pm\cdot0230$ $Q=\cdot3963\pm\cdot0301$
4—5	$r_t=\cdot2428\pm\cdot0212$ $Q=\cdot3434\pm\cdot0304$	$r_t=\cdot3524\pm\cdot0196$ $Q=\cdot4603\pm\cdot0232$	$r_t=\cdot3677\pm\cdot0209$ $Q=\cdot5089\pm\cdot0231$	$r_t=\cdot2978\pm\cdot0246$ $Q=\cdot4306\pm\cdot0301$
6—7	$r_t=\cdot2373\pm\cdot0240$ $Q=\cdot3653\pm\cdot0383$	$r_t=\cdot2777\pm\cdot0230$ $Q=\cdot3963\pm\cdot0301$	$r_t=\cdot2978\pm\cdot0246$ $Q=\cdot4306\pm\cdot0301$	$r_t=\cdot3128\pm\cdot0276$ $Q=\cdot4705\pm\cdot0326$

and Second Brother in terms of First, and the repetition of the same numbers in a second cell does not give those numbers double weight. Had we worked our tables for Elder Brother and Younger Brother, each cell would have had independent weight, but adding them reversed we have lost one-half of the non-diagonal independent frequencies, and we must not still retain the same number of independent weights. Relative to Mr Yule's method of procedure, this, in our opinion, true method of weighting emphasises in symmetrical tables the diagonal columns and would correspondingly *better* tetrachoric r_t as against Q. But we have not used what we consider the true weighting, because it might be said that it had been adopted with a view to bettering our position. Arranging as before we have the following results for the 16 coefficients:

	Weighted Mean	Weighted Standard Deviation
Tetrachoric r_t ...	·3474	·0931
Association Q ...	·5083	·0952

Thus the positions of r_t and Q are just reversed by proper weighting*. The coefficient of variation has no meaning, we hold, in the case of mere numerics like r_t and Q, both of which may range from -1 to $+1$ in the general case. Indeed the case is more complex than can be accurately determined even by weighted standard deviations. For Q, although nominally ranging between ± 1, is for any given case numerically greater than the corresponding tetrachoric r_t, and thus their variabilities if not functionally related are related by limitations. For reasons already given we see no advantage in considering the probable error of

* Mr Yule (*loc. cit.* p. 634) gives them as ·084 and ·081. If we take into account all the reasonably possible divisions 36 in number, i.e. 1–2, 2–3, 3–4, 4–5, 6–7, 7–8, the standard deviations are respectively $\sigma_r=\cdot121$ and $\sigma_Q=\cdot130$, while the ranges, on which Mr Yule appears to lay stress, take for r_t the value: ·59 and for Q: the value ·92!

colligation ω (it is quite easy to deduce another "coefficient of association" from Q, which will have a still less probable error than ω, or one from tetrachoric r_t having a less value than both: see p. 262 ftn.). Further we do not consider that the product-moment method (i.e. ϕ which Mr Yule erroneously terms the correlation coefficient) has any application to these eye-colour data; it is purely idle to deal with the difference between a 'light brown' and 'dark brown' eye as a discrete 'unit.'

We next turn to the Father and Son Eye-Colour data. The table for 16 divisions is:

For Father.

For Son.	*2—3*	*3—4*	*4—5*	*6—7*
2—3	$r_t = \cdot 504 \pm \cdot 029$ $Q = \cdot 616 \pm \cdot 030$	$r_t = \cdot 405 \pm \cdot 031$ $Q = \cdot 528 \pm \cdot 038$	$r_t = \cdot 385 \pm \cdot 035$ $Q = \cdot 549 \pm \cdot 050$	$r_t = \cdot 316 \pm \cdot 041$ $Q = \cdot 487 \pm \cdot 064$
3—4	$r_t = \cdot 391 \pm \cdot 031$ $Q = \cdot 500 \pm \cdot 038$	$r_t = \cdot 550 \pm \cdot 027$ $Q = \cdot 658 \pm \cdot 027$	$r_t = \cdot 493 \pm \cdot 032$ $Q = \cdot 632 \pm \cdot 034$	$r_t = \cdot 421 \pm \cdot 038$ $Q = \cdot 579 \pm \cdot 045$
4—5	$r_t = \cdot 276 \pm \cdot 035$ $Q = \cdot 381 \pm \cdot 049$	$r_t = \cdot 466 \pm \cdot 031$ $Q = \cdot 590 \pm \cdot 034$	$r_t = \cdot 575 \pm \cdot 032$ $Q = \cdot 716 \pm \cdot 028$	$r_t = \cdot 517 \pm \cdot 037$ $Q = \cdot 684 \pm \cdot 035$
6—7	$r_t = \cdot 266 \pm \cdot 040$ $Q = \cdot 402 \pm \cdot 062$	$r_t = \cdot 374 \pm \cdot 038$ $Q = \cdot 519 \pm \cdot 047$	$r_t = \cdot 457 \pm \cdot 040$ $Q = \cdot 622 \pm \cdot 040$	$r_t = \cdot 512 \pm \cdot 039$ $Q = \cdot 695 \pm \cdot 037$

In this case we find:

	Weighted Mean	Weighted Standard Deviation
Tetrachoric r_t ...	·443	·086
Association Q ...	·605	·086

If we thought range was a measure of variation we should have:

Range of tetrachoric r_t = ·31.

Range of association Q = ·34.

It is clear that for these eye-colour tables Q is almost as stable as tetrachoric r_t. Indeed judged by the probable error test Q is slightly better than tetrachoric r_t, for we have

Mean value of	Brothers and Brothers	Fathers and Sons
$(r_t - \bar{r}_t)/(\text{p.e. of } r_t)$ ...	3·5	2·2
$(Q - \bar{Q})/(\text{p.e. of } Q)$ ...	3·2	1·9

The weighted standard deviation test, which is probably a better one, shows that r_t is superior to Q for the Brothers table and equal to it for the Father and Son table; both are terribly bad and a slight change in the nature of the test may make one or other apparently slightly superior. In this eye-colour material in its present state we are quite clear that the tetrachoric method applied to extremely skew divisions will not give consistent results. But this had been stated years ago, in passages which Mr Yule refrains from quoting. We are equally clear that Q gives no better results, while for the bulk of tables of statistical practice, it certainly gives less stable results than r_t.

(G) *Age of Husband and Wife Data.*

This is a case of extreme skewness* selected by Mr Yule to show how tetrachoric r_t varies. It is also a case of heterogeneity, for second marriages are mixed up with first marriages. With his usual ingenuity Mr Yule has exhibited a curve and a table, in which he has increased the number of the terminal values, where tetrachoric r_t gives too low values and has little weight, and shown few of the central values where r_t is relatively steady and has great weight. Further, he tells us that:

"From the standpoint of the calculator, however, the table presents the disadvantage that the correlation is high, viz. ·91, and the approximation to the value of the normal coefficient correspondingly slow, eight to ten or twelve terms of the equation being necessary to give a value fairly trustworthy in the second place of decimals" (*loc. cit.* p. 624). Mr Yule may have got results "fairly trustworthy in the second place of decimals," but as either the odd or even series of powers of r_t in the equation rarely becomes *convergent* till much beyond the twelfth term—we have had in some cases to go to 18 or 20 terms—the confidence Mr Yule put in his values was quite unjustified. The case is an interesting one because the true correlation is very high, i.e. $\cdot 9253 \pm \cdot 00004$†. Accordingly, Mr Yule's association coefficient whatever division is taken is constrained to lie between something like ·952 and 1. It is therefore difficult to compare the range of instability of Q with that of tetrachoric r_t. We are bound to consider both in relation to their possible ranges.

That the curve given for tetrachoric r_t indicates no defect in r_t relatively to other coefficients, will be at once appreciated by comparing it with the curve for Mr Yule's coefficient of colligation ω! That coefficient varies just as much, only it is the other way round: see Diagram XXI, p. 278. As for the Boas-Yulean coefficient ϕ, we could only assume from it, that there was, when the dichotomies were at young or old ages, a low relation between the ages of Husband and Wife.

* The actual skewnesses are: for Husbands Sk. $= \cdot 71$ and for Wives Sk. $= \cdot 76$. These are among the highest values on record for skewness, and the surface is not "moderately skew" as Mr Yule without publishing any numbers (*loc. cit.* p. 624) states it to be.

† The regression line is sensibly curved at the terminals, but this does not markedly influence η which uncorrected is ·9142, as against $r = \cdot 9136$ without Sheppard. With Sheppard r rises to ·9253. Mr Yule gives ·91 for this correlation, which appears to us an uncorrected value.

Yet according to Mr Yule ϕ is "applicable in its entirety to the 2×2-fold table." Of course we may "apply" any method to any problem, but whether we shall obtain anything of value from the application is another question, and Pearson's original statement about Boas' coefficient that "it differs in the simplest cases from the true coefficient of correlation, and often differs considerably...and its use is liable to be misleading, especially if compared with values of the true coefficient found by other processes*" was amply justified and is well illustrated by this case. The tetrachoric r_t values cluster round the true value, the Boas-Yulean never reaches it. In the following table the values of the tetrachoric r_t, of the coefficient of association Q, of the coefficient of colligation ω, and of the Boas-Yulean, Pearson's ϕ, are recorded with their probable errors for all divisions at *like* ages, since these are the divisions selected by Mr Yule.

Ages of Husband and Wife. Values of the various Coefficients proposed to measure Association.

Division of age	r_t	Q	ω	ϕ
18	?	·9991 ± ·00018	·9587 ± ·0004	·0711 ± ·0097
19	?	·9965 ± ·00020	·9197 ± ·0022	·1240 ± ·0042
20	·7755 ± ·0025	·9932 ± ·00016	·8895 ± ·0012	·2128 ± ·0024
25	·8813 ± ·0030	·9734 ± ·00010	·7919 ± ·0003	·5644 ± ·0005
30	·9302 ± ·0001	·9745 ± ·00006	·7958 ± ·0002	·7137 ± ·0002
35	·9522 ± ·0001	·9795 ± ·00005	·8151 ± ·0002	·7735 ± ·0002
40	·9535 ± ·0001	·9821 ± ·00004	·8265 ± ·0002	·7948 ± ·0002
45	·9630 ± ·0001	·9846 ± ·00004	·8381 ± ·0002	·8003 ± ·0002
50	·9601 ± ·0001	·9850 ± ·00004	·8401 ± ·0002	·7867 ± ·0002
55	·9570 ± ·0001	·9863 ± ·00004	·8468 ± ·0002	·7658 ± ·0002
60	·9471 ± ·0001	·9864 ± ·00005	·8471 ± ·0002	·7260 ± ·0003
65	·9350 ± ·0002	·9881 ± ·00005	·8565 ± ·0003	·6733 ± ·0005
70	·9159 ± ·0003	·9897 ± ·00006	·8656 ± ·0003	·5947 ± ·0007
75	·8917 ± ·0006	·9922 ± ·00006	·8823 ± ·0005	·4932 ± ·0012
80	·8450 ± ·0020	·9946 ± ·00008	·9013 ± ·0007	·3512 ± ·0024
85	·8081 ± ·0024	·9975 ± ·00010	·9317 ± ·0013	·2046 ± ·0046

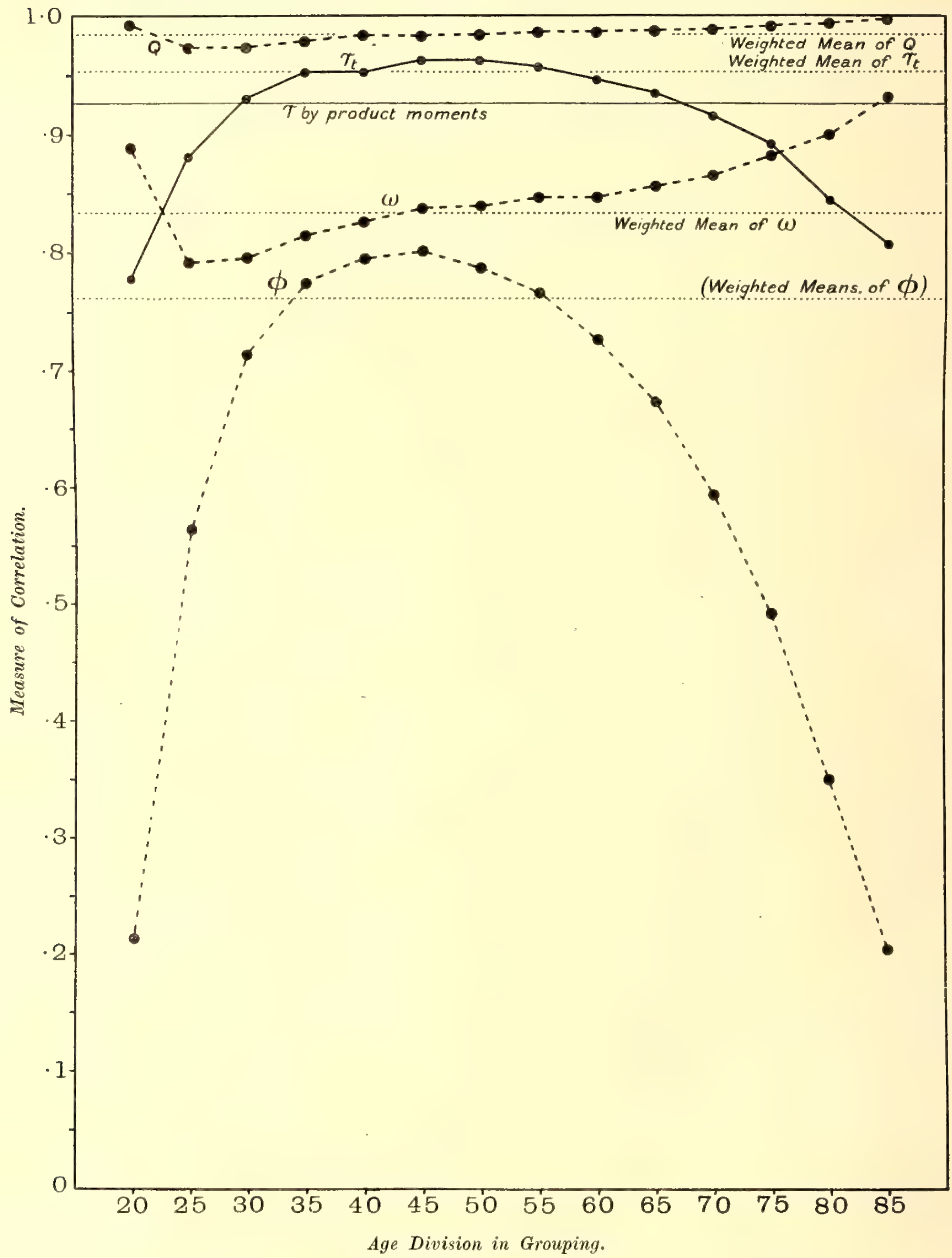
The first two divisions were not used because the work of calculating tetrachoric tables to the large number of terms in r_t required for getting results even approximately correct is excessive when r_t is large and the dichotomies extreme. It will be seen that we very frequently differ by a unit in the second figure of r_t from Mr Yule's results and that our diagram (Diagram XXI) differs from his for r_t by being sensibly nearer to the true correlation. He has stopped his series before it began to converge properly.

The following gives the percentage in the least quadrant at each age division:

18	*19*	*20*	*25*	*30*	*35*	*40*	*45*	*50*	*55*	*60*	*65*	*70*	*75*	*80*	*85*
·0003	·005	·031	·962	1·92	2·13	2·12	1·88	1·69	1·35	1·10	·734	·444	·216	·081	·009

* *Science*, Vol. XXX. p. 24.

DIAGRAM XXI. Diagram showing variations of various coefficients for fourfold tables with diagonal dichotomies in the case of Ages at Marriage of Husband and Wife.



The reader may ask what are our actual grounds for neglecting the results obtained from extreme divisions in the case of tetrachoric r_t? The answer is that it is not only experience of how slight skewness at *extreme* divisions produces large changes in the value of r_t, but our knowledge that the *weights* of these outlying divisions are in the case of r_t (to a lesser extent in the case of ϕ) insignificant as compared with the weights of the central divisions. No such grave differences exist in the case of Q or (to a slightly lesser extent) in the case of ω. The following table gives the weights, treating the weights of the division at 20 as unity.

Division	r_t	Q	ω	ϕ
20	1	1	1	1
25	69	2·7	12·5	24·5
30	370	6·1	27·4	97·9
35	772	10·9	36·2	171·9
40	977	14·3	40·1	187·7
45	1276	16·7	44·7	153·1
50	977	15·8	40·1	145·6
55	625	15·0	31·4	97·1
60	319	11·9	25·2	55·4
65	142	10·0	18·5	22·6
70	54	7·7	12·5	10·6
75	16	5·9	6·9	3·8
80	1·5	3·7	3·0	1·1
85	1·1	2·5	0·9	0·3

It will be clear: first that it is idle to measure the variations of these quantities from anything but a weighted mean, and secondly that no one would after seeing such weights dream of determining tetrachoric r_t from *extreme* divisions. If we omit the first two and the last two values of r_t, noting their slight weight, then the remaining values in no case differ by ·04 from the true correlation and the mean divergence is ·015 only. Mr Yule's Q must be steadier here than r_t, because its range is limited by the nature of the case to about ·05, while there is no limit to the range of the latter.

(H) *Association in a typical table of Ordinary Statistical Practice.*

Thus far we have dealt with the relative stability of the coefficient of association and tetrachoric r_t on material especially selected by Mr Yule to exhibit the variable character of tetrachoric r_t. But in concluding this branch of our discussion we should like to exhibit the relationship of the two coefficients for such surfaces as occur in ordinary statistical practice. For this purpose we will make no selection ourselves, but illustrate the matter on the correlation table chosen by Mr Yule himself in his first memoir on Association*. We take the table as it is given without any knowledge of its degree of approach to the Gaussian or of the material with which it deals. Its coefficient of correlation by the product-moment method = ·677. Now we commenced by leaving out all the

* See *Phil. Trans.* Vol. 194 A, p. 277.

Association Q and Tetrachoric r_t with Probable Errors for Mr Yule's Table.

	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	— —	·954 ± ·026 ·674 ± ·101	— —	·899 ± ·053 ·578 ± ·095	— —	— —	— —	— —	— —	— —	— —	— —	— —	— —
3	·954 ± ·023 ·700 ± ·084	— —	·905 ± ·028 ·658 ± ·059	— —	·792 ± ·054 ·517 ± ·061	— —	·700 ± ·087 ·414 ± ·067	— —	— —	— —	— —	— —	— —	— —
4	— —	·914 ± ·031 ·657 ± ·064	— —	·861 ± ·024 ·671 ± ·038	— —	·817 ± ·032 ·615 ± ·039	— —	·888 ± ·037 ·627 ± ·036	— —	— —	— —	— —	— —	— —
5	·893 ± ·063 ·570 ± ·090	— —	·842 ± ·034 ·616 ± ·049	— —	·802 ± ·023 ·667 ± ·031	— —	·837 ± ·021 ·694 ± ·025	— —	·918 ± ·024 ·699 ± ·025	— —	— —	— —	— —	— —
6	— —	·917 ± ·056 ·608 ± ·080	— —	·846 ± ·031 ·634 ± ·027	— —	·786 ± ·021 ·669 ± ·024	— —	·811 ± ·019 ·706 ± ·021	— —	·882 ± ·024 ·687 ± ·025	— —	— —	— —	— —
7	— —	— —	·855 ± ·062 ·526 ± ·048	— —	·861 ± ·031 ·657 ± ·028	— —	·789 ± ·021 ·674 ± ·022	— —	·808 ± ·018 ·706 ± ·022	— —	·866 ± ·026 ·673 ± ·030	— —	— —	— —
8	— —	— —	— —	— —	— —	·880 ± ·030 ·652 ± ·027	— —	·832 ± ·021 ·689 ± ·024	— —	·831 ± ·019 ·711 ± ·026	— —	·926 ± ·019 ·747 ± ·033	— —	— —
9	— —	— —	— —	— —	— —	— —	·913 ± ·031 ·663 ± ·031	— —	·888 ± ·021 ·716 ± ·029	— —	·845 ± ·023 ·673 ± ·040	— —	·949 ± ·018 ·751 ± ·046	— —
10	— —	— —	— —	— —	— —	— —	— —	·927 ± ·034 ·660 ± ·039	— —	·900 ± ·025 ·696 ± ·042	— —	·897 ± ·023 ·690 ± ·045	— —	·864 ± ·064 ·519 ± ·100
11	— —	— —	— —	— —	— —	— —	— —	— —	— —	— —	·867 ± ·046 ·577 ± ·070	— —	·889 ± ·048 ·560 ± ·095	— —
12	— —	— —	— —	— —	— —	— —	— —	— —	— —	— —	— —	·793 ± ·103 ·427 ± ·116	— —	·829 ± ·101 ·468 ± ·119

The first figures in each cell refer to Q, the second to r_t.

border cases of the table which give $Q = \pm 1$ with infinite weight. Also the cases adjacent to these where Q, depending upon very small frequency in one quadrant, has a very high value, a low probable error, and accordingly a very great weight. With the exception of this we took every alternate case in every row and so obtained 41 coefficients; the corresponding tetrachoric r_t's were likewise calculated. The results are exhibited in the accompanying table, the notation being that adopted by Mr Yule to mark the divisions.

There is practical certainty that any extension of the boundaries of this system of divisions could only better the position of r_t relative to Q. We obtained the following results:

$$\text{Weighted Mean } Q = \cdot 834. \qquad \text{Weighted Mean } r_t = \cdot 672.$$

$$\text{Weighted S.D. of } Q = \cdot 0613. \qquad \text{Weighted S.D. of } r_t = \cdot 0478.$$

We must therefore increase the variability of tetrachoric r_t by 28 °/₀ to reach the variability of Q. This shows the real degree of difference in the stabilities of r_t and Q for the correlation-tables of ordinary practice. We note also that the weighted mean of r_t differs by only ·005 from the true product-moment value, ·677, of the correlation.

Now let us approach the matter from the standpoint of probable error. We have:

$$\text{Mean } (r_t - \bar{r}_t)/\text{probable error} = 1\cdot 06.$$

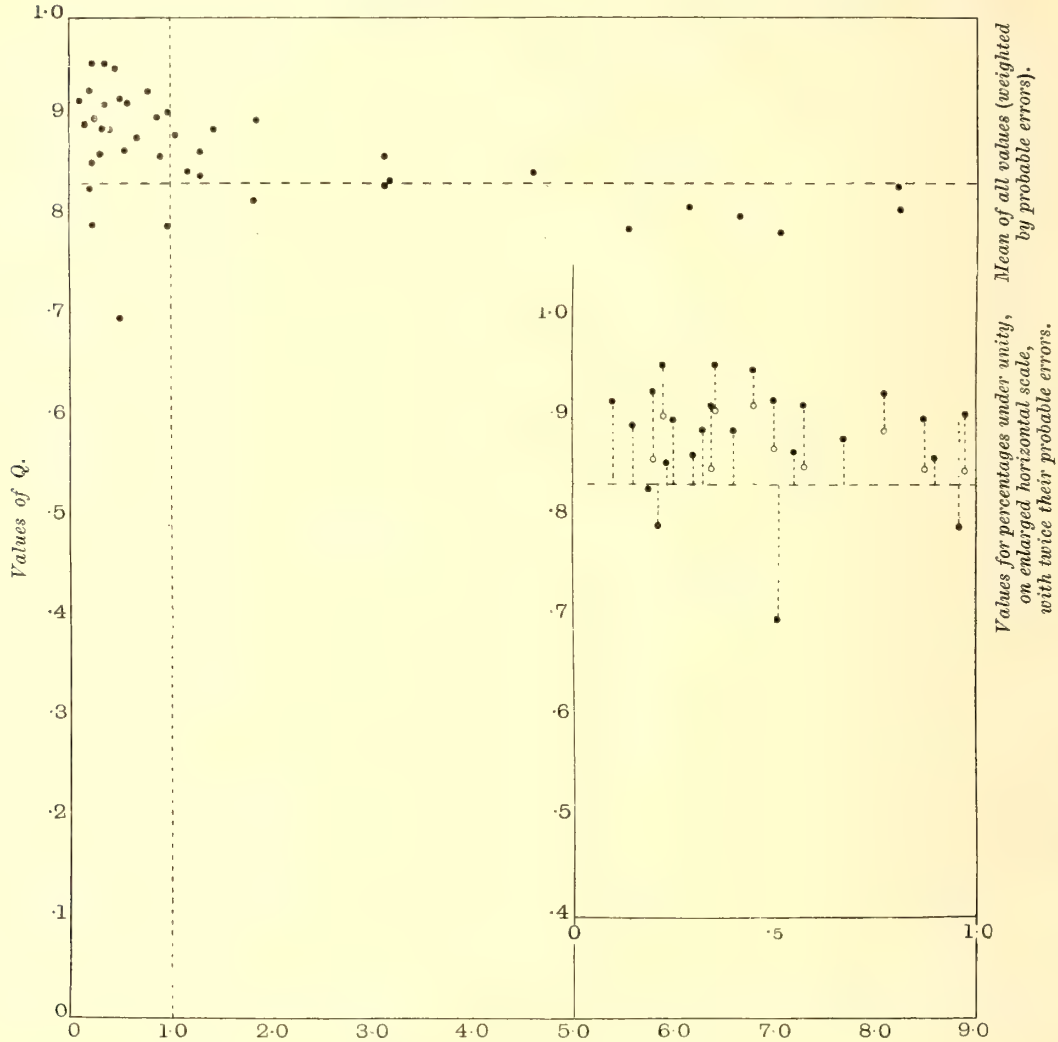
$$\text{Mean } (Q - \bar{Q})/\text{probable error} = 1\cdot 72.$$

In the case of $r_t - \bar{r}_t$ we have 21 below their probable error in value and 20 above it, just what there should be. Only 5 exceed twice the probable error and there should be 4. In other words the distribution of r_t in terms of its probable error might well have arisen from random sampling of Gaussian material. Now turn to Q: in terms of the probable error Q shows an increase of upwards of 62 °/₀ on the value for r_t. There are only 15 values of $Q - \bar{Q}$ below this probable error, 26 in excess of it. There are 15 values instead of 4 in excess of *twice* the probable error. There are five values in excess of three times their probable error, compared with only two occurring in the case of r_t. It is, we think, obvious that the variations in the case of Q are far greater than those due to random sampling*. Our Diagrams XXII and XXIII (pp. 282–3) indicate two points. In the upper parts of these diagrams we have plotted r_t and Q to the percentage of frequency in the quadrant of least frequency. We see at once that if we avoid quadrantal frequencies under 1 °/₀ the value of tetrachoric r_t is for practical purposes equal to the true product-moment r. The reader will recognise the far greater scatter of Q. In the lower figures we have plotted r_t and Q with relation to their probable errors; the full dot denotes the observation, the open dot the end of a line

* The reader must remember that there is no reason to assert that the errors must be of the order of random sampling; there is only one table, not many random samplings from a much larger mass, and we take different divisions. But we assert that if the errors be of that order, then the method is as good as the data warrant our using.

equal to twice the probable error; where this open dot is seen, there twice the probable error fails to reach the true value of r, or the mean value of Q as the case may be. The failures of Q are seen to be three times as many as those of r_t.

DIAGRAM XXII. Values of association Q for a typical table of ordinary statistical practice.

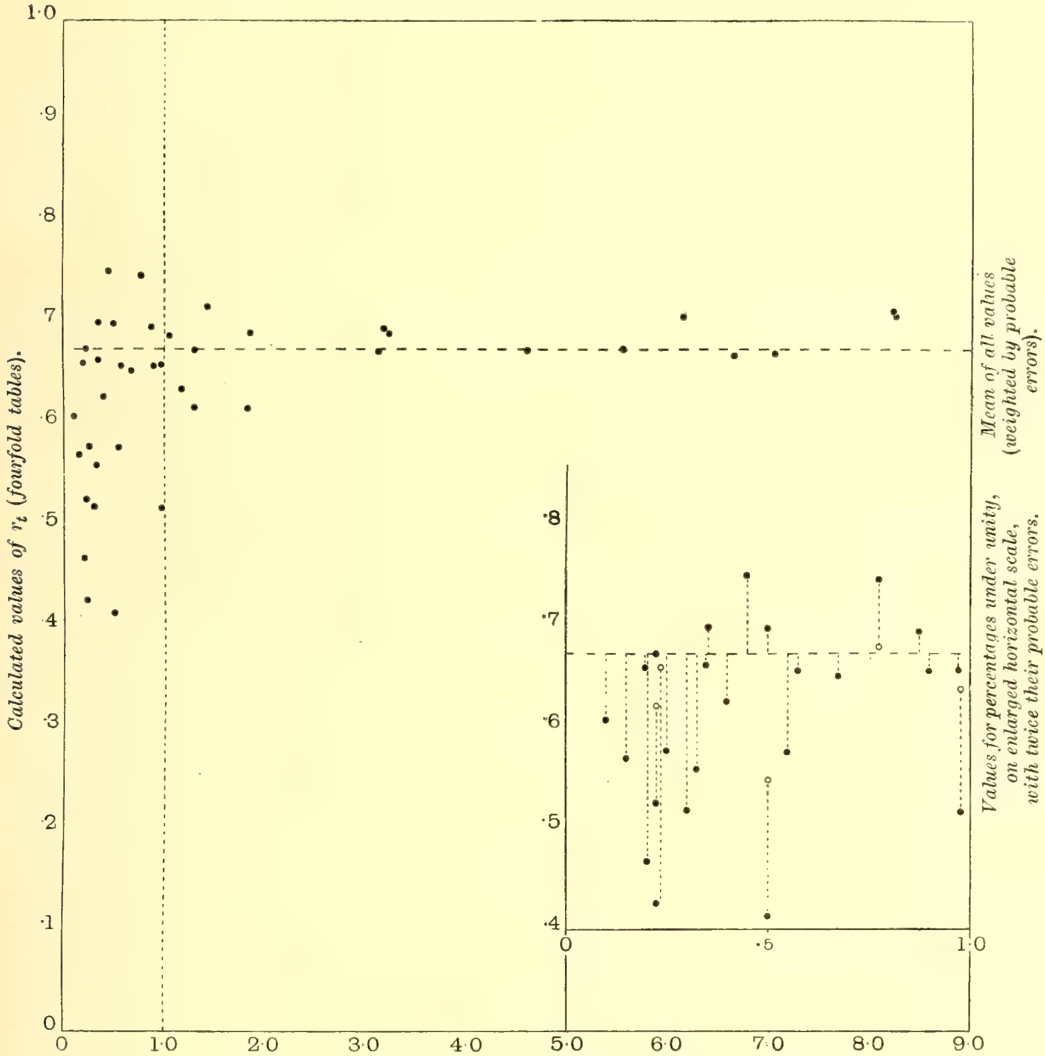


We may conclude justifiably that even as a coefficient of association r_t is much more stable than Q for the usual type of distribution; and we have already seen that this is so even for many marked cases of skewness. It is always an approximation and often a very good approximation to the true correlation. On the other hand Q conveys no idea to the mind at all, except in as far as it is an approximation, usually very bad, to the same correlation.

(14) *General Protest against the Use of Mr Yule's Coefficient of Association.*

It may be said that a vigorous protest against Mr Yule's coefficients is unnecessary. We believe on the contrary that, if not made now and made strongly, there will be a great set-back to both modern statistical theory and

DIAGRAM XXIII. Values of r_t for a typical table of ordinary statistical practice.



Percentage of total in smallest quadrant.

modern statistical practice. The publication of Mr Yule's text-book has resuscitated the use of his coefficient of association; it is now being used in all sorts of quarters for all sorts of unsuitable data. The publication of Mr Yule's recent paper on Association will also lead to the use of his method of pseudo-ranks. The

coefficients of association and colligation are in our opinion wholly fallacious, they represent no true properties of the actual distributions, and they have no adequate physical interpretation. The coefficient obtained by the method of pseudo-ranks is equally fallacious, unless the variables proceed by and have been grouped by discrete units. But both Mr Yule's methods are so easy of application, that those who will not devote the small amount of time and energy requisite to the discussion of data by more adequate processes at once adopt them without further consideration. Thus Mr Yule's coefficient of association is passing into French statistical literature as *l'indice de corrélation*, a term originally introduced by Galton for the coefficient of correlation and now transferred to the different and we hold fallacious Yulean measure of association. Professor Niceforo writes *: "Mais pour nos études sur la corrélation entre les différents phénomènes économiques, démographiques et autres, dans les quartiers et les arrondissements des grandes villes, où nos trouvions en présence de séries formées par un nombre plutôt restreint d'éléments (80, 25, 20) nous avons préféré nous servir de la méthode Yule, plus rapide et donnant, quoique moins précise que la méthode précédente, des résultats très satisfaisants." The preceding method is the method of the product-moment which Professor Niceforo discards for Mr Yule's association coefficient, using it solely—and apparently with Mr Yule's approval (see *loc. cit.* p. 324)—for *absolutely continuous* variates, where the coefficient of correlation could be at once found and a graph easily drawn of the regression line. Professor Niceforo speaks everywhere in his paper of the correlation being this or that, and entitles his paper "Contribution à l'étude des *corrélations* entre le bien-être économique et quelques faits de la vie démographique." What his or Mr Yule's test of "résultats très satisfaisants" may be we do not know, but we consider that the whole of Professor Niceforo's work will have to be repeated before anything can be learnt from his data.

We have worked out two illustrative cases from Professor Niceforo's material to indicate what we consider the extreme danger of Mr Yule's methods. In the first case, that of the average stature of conscripts and the average rent in the 20 arrondissements of Paris, the answer Professor Niceforo gives is $Q = 1 \pm \cdot 0$. The correlation we are told is perfect. In the second case, that of the correlation between the mortality and natality of the same arrondissements, we are told that the "corrélation...est très forte." Professor Niceforo gives it as:

$$\text{Indice de corrélation } R = 0\cdot 977 \pm \cdot 032{,}450 \ (\textit{sic}\,!).$$

We had failed to give any real interpretation to either of these results, and we turned to the original data. These we found hard to discover, because Professor Niceforo does not refer at each stage to the exact source of his original material. Manouvier gives the mean stature of the 20 arrondissements in 1880 and 1881†. These appear to be what is given as *Statura media* in Professor Niceforo's book

* *Journal de la Société de Statistique de Paris*, 52 Année, pp. 322—341; see p. 324 and elsewhere.

† *Bulletins de la Société d'Anthropologie de Paris*, Année 1888, p. 161.

Forza e Ricchezza, Turin, 1906, p. 15. For the rents we have not been able to verify Professor Niceforo's returns, or to discover whether they are for approximately the same years as the statures. In the *Annuaire Statistique de la Ville de Paris*, 1901, no similar details as to rent appear to be given, and Professor Niceforo gives no reference to the year for his rent data. Taking, however, the data given in his book we find:

Mean Stature = 1645·75. Mean Rent* = 210·8.

Standard Deviation = 6·212. Standard Deviation = 161·43.

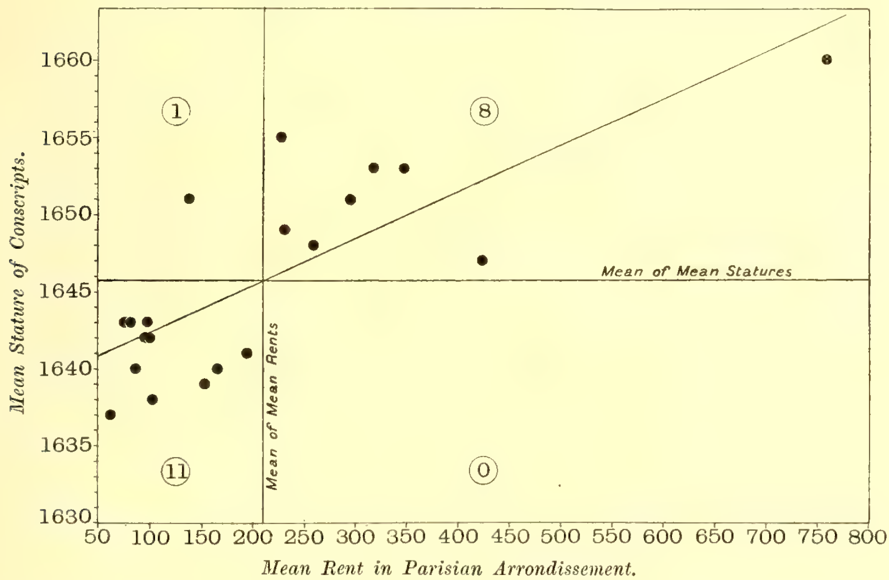
Correlation of Stature and Rent = ·7825 ± ·0591.

Regression line of Stature S on Rent R:

$$S = 1639{\cdot}4 + {\cdot}03011\, R.$$

This is represented on Diagram XXIV and we see at once that we have a quite intelligible relation between average rent in the arrondissement and average stature of conscripts. The correlation is high, but far from "perfect," and is subject

DIAGRAM XXIV. Association of Stature and Rent in Parisian Arrondissements.



to a large probable error, so that the true relation might be easily anything between ·66 and ·90. If, however, we divide at the mean values and add up all the dots as Professor Niceforo appears to have done, we have the fourfold table: $\begin{array}{c|c} 1 & 8 \\ \hline 11 & 0 \end{array}$ giving association $Q = 1 \pm {\cdot}0$. What is the value of this, what does it signify? We fail to extract the least idea of the real relationship—as represented by the dots on the graph—from such a statement. *Verbally* it means simply that no rent of over 210·8 is associated with a stature under 1645·75, but to assert that

* We are unable to say in what units or for what periods rent is measured.

this signifies that "la corrélation est parfaite" is totally misleading. Professor Niceforo speaks of the correlation between stature of conscripts and the rent of their arrondissements of origin; he does not say that he is merely dealing with the practically unimportant fact that no arrondissement with a mean rent over 211 has given a mean stature under 1646. But even to state this simple fact would be more enlightening than to talk of the index of correlation equalling unity.

The second case we took was that of the mortality and natality of the same 20 arrondissements. Again we had great difficulty in tracing the original source of the information.

However in the *Annuaire Statistique de la Ville de Paris*, Année 1904, we find, p. 125, the natality per year based upon 1000 women of ages 15—49, and, on p. 135, the annual mortality based on 1000 inhabitants for these arrondissements. The natality is for the period 1886—1895, the mortality is given for the periods 1886—1890 and 1891—1895 separately but not combined. As Professor Niceforo does not give a reference to the years dealt with, nor the source of this Parisian data, we have taken the simple mean of the mortality for the two periods and correlated this with the natality for 1886—1895. The constants found are:

Mean Mortality $= 21{\cdot}73$. Mean Natality $= 79{\cdot}065$.

Standard Deviation $= 4{\cdot}9215$. Standard Deviation $= 23{\cdot}7298$.

$$\text{Correlation} = r = {\cdot}9163 \pm {\cdot}0242.$$

Regression Line of Natality N on Mortality M:

$$N = 4{\cdot}4181\, M - 16{\cdot}94.$$

Diagram XXV indicates the position of the observations and their relation to the regression line. It conveys as adequate a representation of the whole relationship as it is possible to give on the data. But if we count up on the graph the dots in the quadrants obtained by drawing the lines at the means we obtain the fourfold division $\begin{array}{c|c} 0 & 8 \\ \hline 11 & 1 \end{array}$ leading again to $Q = 1 \pm {\cdot}0$*. What information as to the real nature of the correlation is given by such a result?

If Professor Niceforo had desired to obtain a rapid approximate value to the true correlation in these cases, he should have drawn his divisions at the medians, not the means, and used Sheppard's formula

$$r = \cos \pi \frac{b}{a+b} \quad \text{for} \quad \begin{array}{c|c} a & b \\ \hline b & a \end{array}$$

to find tetrachoric r_t. In this case his results would have been the tables

$$\begin{array}{c|c} 8 & 2 \\ \hline 2 & 8 \end{array} \quad \text{and} \quad \begin{array}{c|c} 9 & 1 \\ \hline 1 & 9 \end{array}$$

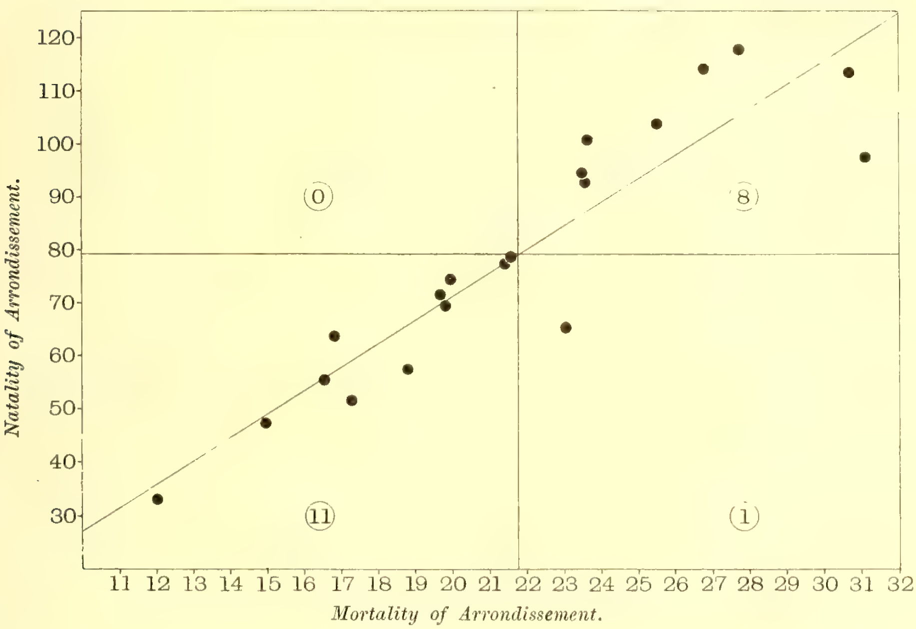
$$r_t = {\cdot}81 \pm {\cdot}05, \text{ for rent and stature,}$$

$$r_t = {\cdot}95 \pm {\cdot}04, \text{ for natality and mortality.}$$

* Professor Niceforo gives ·977 probably from slightly different data.

Both results correspond within the limits of the probable errors with the actual correlations ·78 and ·92. These results are far more valid than the association results, but of course have not the value of a graph showing the regression line. Now here is a case of a man proposing to deal with a most interesting problem, for which quite serviceable data exist, led at once from the track of sound treatment by the application of this fallacious doctrine of association! And this is far from a solitary instance of the harm Mr Yule has done by the publication without adequate warning or guidance to his readers of the section of his text-book treating of association.

DIAGRAM XXV. Association of Natality and Mortality in Parisian Arrondissements.



(15) *Further Criticisms of Mr Yule's Methods of Controversy.*

To certain minor points of Mr Yule's memoir reply will be made in the present section.

(*a*) *Partial Correlations formed from the Normal Coefficient* (*loc. cit.* p. 627). Mr Yule is deliberately confusing two different ideas, the correlation of A and B, two continuous variables, for a *constant value* of a third variable C, with the correlation of A and B for a given range of values summed under a certain class-index or group of class-indices of the variable C. The former is the only sense in which the term partial correlation has been hitherto used, and there is no reason why Mr Yule should deliberately confuse this sense with a wholly different conception, that of the correlation of A and B for a sub-universe of C. The whole

theory of the correlation of sub-universes has been dealt with by Pearson* and the formulae obtained in 1901 were known shortly afterwards to be perfectly general, although the proof of this generality was only recently published†.

Let s_1 be the standard deviation of the sub-universe selected, σ_1 its standard deviation before selection. Let the three characters be represented by the subscripts 1, 2, 3; then if we write

$$\rho_{12} = r_{12}\sqrt{1 - s_1^2/\sigma_1^2}, \qquad \rho_{13} = r_{13}\sqrt{1 - s_1^2/\sigma_1^2},$$

we have for the correlation within the sub-universe

$$_1\mathbf{r}_{23} = \frac{r_{23} - \rho_{12}\rho_{13}}{\sqrt{1 - \rho_{12}^2}\sqrt{1 - \rho_{13}^2}} \ddagger$$

and

$$_1r_{23} = \frac{r_{23} - r_{12}r_{13}}{\sqrt{1 - r_{12}^2}\sqrt{1 - r_{13}^2}}$$

for the partial correlation coefficient.

In the case of normal correlation to which Mr Yule is referring, s_1 is the standard deviation of the truncated portion of a normal curve. For his special cases when that portion is one-half the frequency curve

$$s_1^2 = \sigma_1^2 - \bar{x}^2$$

and $\bar{x} = \sqrt{\frac{2}{\pi}}\,\sigma_1$. Hence

$$\rho_{12} = \sqrt{\frac{2}{\pi}}\,r_{12}, \qquad \rho_{13} = \sqrt{\frac{2}{\pi}}\,r_{13},$$

$$_1\mathbf{r}_{23} = \frac{r_{23} - \frac{2}{\pi}r_{12}r_{13}}{\sqrt{1 - \frac{2}{\pi}r_{12}^2}\sqrt{1 - \frac{2}{\pi}r_{13}^2}}.$$

But it is equally feasible to get almost in a line the value of $_1\mathbf{r}_{23}$ for *any* truncated portion of the normal curve other than one-half, and tables for determining the values of $_x\mu_2$ and $_x\mu_1$, the moments of the tail about the severance ordinate at x, giving

$$s_1^2 = \sigma_1^2\,(_x\mu_2 - {_x\mu_1}^2),$$

were calculated by Dr Alice Lee and published in 1908§. These functions were termed the incomplete normal moment functions. Had Mr Yule paid attention to any of this work, he would hardly have published his special illustration and remarked "At present I have not been able to carry the matter further" (*loc. cit.* p. 628). The general formula had been given eleven years ago, and tables from which it was quite easy to calculate special cases were published four years ago!

* *Phil. Trans.* Vol. 200 A, pp. 1—66.

† *Biometrika*, Vol. VIII. p. 437.

‡ *Phil. Trans. loc. cit.* Eqn. (lvi), p. 25.

§ *Biometrika*, Vol. VI. p. 66.

Pearson did not give a name to his coefficient $_1\mathfrak{r}_{23}$ above, but he carefully distinguished it from the partial coefficient $_1r_{23}$ and stated that the former generalised form which did not select at a given value but *round* it was more important for both natural and artificial selection (*loc. cit.* p. 31). Mr Yule has apparently just awoke to the importance of $_1\mathfrak{r}_{23}$, but that is no reason why he should confuse it in the minds of his readers with $_1r_{23}$, or lead his readers to believe that we do not know the difference between the two. To avoid confusion of this kind in future we shall henceforth speak of $_1r_{23}$ as a singular partial correlation and $_1\mathfrak{r}_{23}$ as a plural partial correlation. For, the former expresses the relation between A and B for a single value of C and the latter for a plurality or universe of values of C. In actual practice there is little difficulty in determining $_1\mathfrak{r}_{23}$ if there is enough material, for all we have to do is to take the given universe out of C and correlate the resulting A's and B's. On the other hand, when we speak of the relation of health of child to health of mother for constant employment of mother or constant habits of mother, we do not look upon the universe of employed mothers as a whole or the universe of mothers with bad habits as a whole. We are thinking of employment of mother as a graduated character and parental habit also as a graduated character, and we properly use $_1r_{23}$ to measure the relationship of health in mother and child for a constant grade of employment or constant grade of bad habit in the mother. In this case the use of $_1r_{23}$ has precisely the same justification, if r_{13}, r_{23} and r_{12} are found by tetrachoric tables, as if they had been found by product-moments, *provided the assumption of a Gaussian distribution be reasonably justified for the material in question.* There is no other source of error in the use of $_1r_{23}$ as Mr Yule obscurely seems to indicate. It did not need Mr Yule's numerical illustration (*loc. cit.* p. 629) to prove that $_1\mathfrak{r}_{23}$ for the two sections of an unequally divided normal curve—'defectives' and 'undefectives' (*sic*!)—is in neither case equal to $_1r_{23}$. The two coefficients have different values and different significance whether the frequency be Gaussian or non-Gaussian.

(*b*) *Mr Yule's Failure to distinguish between Criticism of Method and Criticism of Conclusion.*

We have seen in the course of this paper that Mr Yule's coefficient of association automatically rises in all cases examined when our dichotomy is very one-sided. This is very obvious even in skew distributions; compare Diagrams XVIII, XIX, XXI and XXII where the rapid increase of Q for small percentages is obvious. Heron working on the Gaussian surface had demonstrated that this was an absolute necessity which flowed from the theory and that therefore it must follow, even for surfaces only approximately Gaussian, that two or more values of Q were quite incomparable if the dichotomic lines were at different percentages of marginal frequencies. He argued that no valid proof could therefore be based on the relative sizes of Q in a series of tables for which

we knew nothing about the frequency and for which the dichotomic lines gave marginal frequencies having any continuously ascending order.

Such invalid discussions as to the "apparent law that associations were on the whole higher where populations were healthier or more defective" and on the relation of change of association between two defects with age had actually been attempted by Mr Yule*. Heron criticised, it will be seen, not the laws, if they be laws, but Mr Yule's attempt to investigate them by "association." He wrote†: "For precisely similar reasons his discussion of the change of association with age must be dismissed as entirely fallacious. There may be, and probably is, some decrease of association with advancing age, but the enunciation of such a law on the basis of a number of coefficients of association is purely idle." Heron then proceeded to show that the proportions of blind and of mentally deranged both increased steadily from infancy to age, and this signifies that if the frequency surfaces were of the same type and really had the same correlation at each age the Q would steadily decrease with age. It will be seen that Heron's criticism applied solely to Mr Yule's methods; it was an unanswered and, we believe, absolutely unanswerable criticism of the absurdity of trying to deduce laws from Q. How does Mr Yule meet it? He writes‡: "Dr Heron also objects to my conclusion that association decreases with age. His objection appears to be that the product sum correlation does not decrease so markedly or regularly with age in one of my cases that he examined...and that no evidence has been given that the normal coefficient decreases."

The destructive criticism that Q for the Gaussian and for all surfaces of which we have any practical experience increases the more one-sided is the dichotomy—and Mr Yule thinks nothing of ·02 °/。 dichotomies—is not met at all. The criticism was of the *method* of forming an inference, and not as to whether the inference led to a law which could be otherwise substantiated. "There may be, and probably is, some decrease of association with advancing age," wrote Heron, "but the enunciation of such a law on the basis of a number of coefficients of association is purely idle." The truth of the law or its falsity is of no great importance, but that Mr Yule should reach it by a fallacious method is of fundamental importance. Mr Yule seeks by the words "Dr Heron also objects to my conclusion that association decreases with age"—an objection never raised—to confuse the really destructive criticism, that Q, unlike tetrachoric r_t, having no intelligible correction for the one-sidedness of its dichotomies is a function of the dichotomic percentages and therefore two Q's based on different percentages are wholly incomparable.

* *Phil. Trans.* Vol. 194 A, pp. 309 *et seq.*

† *Biometrika*, Vol. VIII. p. 119.

‡ *Loc. cit.* p. 637.

(c) *Fallacies involved in the use of Percentages. Coefficient of Colligation.*

Mr Yule has endeavoured to give his Q a physical meaning by deducing from it for the "equivalent symmetrical table" his coefficient of colligation ω. For such a table

	A	Not-A
B ...	$\sqrt{ad}$	$\sqrt{bc}$
Not-B ...	$\sqrt{bc}$	$\sqrt{ad}$
Totals	$\sqrt{ad}+\sqrt{bc}$	$\sqrt{ad}+\sqrt{bc}$

$$\omega = \frac{\sqrt{ad}-\sqrt{bc}}{\sqrt{ad}+\sqrt{bc}} = \frac{1}{100}\left\{\begin{pmatrix}\text{percentage of } A\text{'s} \\ \text{which are } B\text{'s}\end{pmatrix} - \begin{pmatrix}\text{percentage of not-}A\text{'s} \\ \text{which are } B\text{'s}\end{pmatrix}\right\}.$$

We have given grave reasons for doubting the process by which Mr Yule deduces this table from his original data. But this very method of percentages itself is liable to gross misinterpretation, and illustrations of this occur throughout Mr Yule's text-book*.

Given a fourfold table:

	A	Not-A	
B ...	a	b	$a+b$
Not-B ...	c	d	$c+d$
	$a+c$	$b+d$	N

the percentage difference is $q_v = 100\left(\frac{a}{a+c} - \frac{b}{b+d}\right)$ for the vertical treatment and $q_h = 100\left(\frac{a}{a+b} - \frac{c}{c+d}\right)$ for the horizontal treatment. Which is to be taken as a real measure of the relationship? Mr Yule in his text-book uses either and apparently has some personal scale of values. He gives no probable errors which alone could give any soundness to his discussions. Actually we find

$$\text{P.E. of } q_v = 67{\cdot}449\sqrt{\frac{ac}{(a+c)^3} + \frac{bd}{(b+d)^3}},$$

$$\text{P.E. of } q_h = 67{\cdot}449\sqrt{\frac{ab}{(a+b)^3} + \frac{cd}{(c+d)^3}}.$$

* *Theory of Statistics*, 1911.

Now here are a few of Mr Yule's percentage coefficients:

(1) Imbeciles and Deaf-mutes (Yule, *loc. cit.* pp. 33—34).

	Imbecile	Non-Imbecile	Totals
Deaf-Mute ...	451	14,795	15,246
Non-Deaf-Mute ...	48,431	32,464,323	32,512,754
Totals	48,882	32,479,118	32,528,000

$q_v = 0 \cdot 877$
$q_h = 2 \cdot 809$

(2) Datura (Yule, *loc. cit.* p. 37).

	Violet	White	Totals
Prickly ...	47	21	68
Smooth ...	12	3	15
Totals ...	59	24	83

$q_v = -\ 7 \cdot 840$
$q_h = -10 \cdot 882$

(3) Houses in course of erection in Urban and Rural Districts (Yule, *loc. cit.* p. 62).

	Built	Building	Totals
Urban ...	4960	50	5010
Rural ...	1749	12	1761
Totals ...	6709	62	6771

$q_v = 6 \cdot 714$
$q_h = 0 \cdot 317$

(4) Eye-colour in Father and Son (Yule, *loc. cit.* p. 34).

Father.

Son.	Light	Not-Light	Totals
Light ...	471	148	619
Not-Light ...	151	230	381
Totals ...	622	378	1000

$q_v = 36 \cdot 570$
$q_h = 36 \cdot 457$

(5) Developmental Defects and Dullness (Yule, *loc. cit.* p. 45, where the numbers are reduced to 10,000).

	With Defects	Without	Totals
Dull ...	888	1186	2074
Not-Dull ...	1420	22793	24213
Totals ...	2308	23979	26287

$q_v = 33 \cdot 528$
$q_h = 36 \cdot 951$

These instances will suffice, and now let us sum them in a table with their probable errors.

Index Number	Values of q	Probable Error	Mr Yule's Judgment on the Table
(3), q_h	0·317	±0·163	"Distinct positive Association"
(1), q_v	0·877	±0·029	"High degree of Association"
(1), q_h	2·809	±0·093	"High degree of Association"
(3), q_v	6·714	±3·40	"Distinct positive Association"
(2), q_v	7·840	±5·77	"No Association"
(2), q_h	10·882	±7·93	"No Association"
(5), q_v	33·528	±0·68	"Very high indeed"
(4), q_h	36·457	±2·05	"Shows the tendency to resemblance"
(4), q_v	36·570	±2·05	"Shows the tendency to resemblance"
(5), q_h	36·951	±0·73	"Very high indeed"

Now Mr Yule has used the method of percentages in a curious manner; sometimes he compares $a/(a+c)$ with $b/(b+d)$ but at other times with $(a+b)/N$; sometimes he uses the percentages found both ways, sometimes only found one way. He has throughout failed to give the probable errors of the differences of the percentages, which might have influenced his judgment, but he leaves his readers to believe that some inference as to the intensity of association can be founded merely upon such relative percentage differences. He indeed tells us (*loc. cit.* p. 651) to distinguish between the intensity of an association and the reliability of that intensity, so that we must presume that in speaking of the grade of the association, he does not form his judgment in relation to the probable error. Now in this table we find percentage differences of 7·8 and 10·9 belong to tables which in Mr Yule's judgment exhibit no association, but tables with differences of 0·3 have "distinct positive association" and of 0·8 have "high degree of association." One table with a difference of 36 merely shows the "tendency" to association; another with the same percentage difference has association "very high indeed." For any given table there are six ways in which the difference of percentages can be enumerated, namely

$$\frac{a}{a+c}-\frac{b}{b+d},\qquad \frac{a}{a+c}-\frac{a+b}{N},\qquad \frac{b}{b+d}-\frac{a+b}{N},$$

$$\frac{a}{a+b}-\frac{c}{c+d},\qquad \frac{a}{a+b}-\frac{a+c}{N},\qquad \frac{c}{c+d}-\frac{a+c}{N}.$$

Mr Yule sometimes uses one, sometimes another of these methods to reach his judgment of the degree of association in a table. He has given us his judgment with regard to the association of developmental defects and dullness for the partial universe of those without nerve signs (*loc. cit.* pp. 45—46), he says the association is "very high indeed." It may according to the percentage difference chosen be either 2·92 or 36·96. He has further given us his judgment on the association of developmental defects and dullness for those with nerve signs, he

says it is "very small." It may according to the percentage difference be 5·15 or 0·96; thus the percentage difference might be higher than in the very "high association indeed" case in the partial universe of those without nerve signs.

Finally in the "high association" case of the total universe, the percentage difference might be 1·91 or 51·03. Clearly in judging by percentages the conclusion will depend on which percentage is worked out first. Is it not clear that, however Mr Yule may have reached his judgments of no, small, high or very high association, the percentage difference is not his actual measure and could only confuse the tyro in statistics, for whom he introduces this "simple" method? But if difference of percentage has obviously no correlation with Mr Yule's judgment of these grades of association, what weight can possibly be given to the coefficient of colligation in determining association? The chief merit of that coefficient according to Mr Yule is that it has—what Q has not—a physical meaning; but for him percentage differences of 1·91 and 36·96 alike mark "very high associations" and differences of 0·96 and 10·88 alike mark very small or no association. We have no standard and clearly Mr Yule has none of how such differences of percentages are to be interpreted. Mr Yule has by his treatment of percentages *a priori* destroyed any rational meaning that could be given to his own coefficient of colligation as a measure of relationship. Is not Professor Edgeworth's question answered? The coefficient is a "colligation," not a "profound truth." Mr Yule obviously lays no consistent stress whatever on percentage differences in practice.

(*d*) *Mr Yule's use of Pearson's "Transfer"* $(ab-cd)/N$ *as a Measure of Association.*

Among Mr Yule's many means of testing association—no two of which give as a rule the same result—perhaps the most striking is his use of the "Transfer" $(ab-cd)/N$, to which he gives a new name and letter, the "common difference" δ*. One of us had already suggested that the "transfer" *per unit of the total frequency* might be used as a coefficient of association of an inadequate character†. It is inadequate because it makes no correction for class-indices and none for the centroids of the quadrants. Thus it does not lie between 0 and 1, and is largely affected by the position of the dichotomic lines. Mr Yule has preferred to use instead of the transfer per unit of total frequency simply the transfer, his "common difference δ," and the results add further evidence of the vagueness of the whole of his conceptions of association. He tells us, to begin with, that "the difference of the cross-products may be very large if N be large, although δ is really very small... the difference should be compared with N, or it will be liable to suggest a higher degree of association than actually exists" (*loc. cit.* p. 37). To illustrate his

* *Theory of Statistics*, p. 36.

† *Phil. Trans.* Vol. 195 A, p. 14.

method he uses some of Bateson and Saunders' data for *Datura* from the *Report to the Evolution Committee*, 1902:

Colour of Flower.

Fruit.	Violet	White	Totals
Prickly ...	47	21	68
Smooth ...	12	3	15
Totals ...	59	24	83

The difference of the cross products is $252 - 141 = 111$ and then Mr Yule proceeds to tell us that "at first sight this considerable difference is apt to suggest a considerable association." He then divides by 83, and writes: "But $\delta = 111/83 = 1 \cdot 3$ only, so that in point of fact the association is small, so small that no stress can be laid on it as indicating anything but a fluctuation of sampling" (p. 37).

That Mr Yule is content with this process is evident from the opening words of the following paragraph: "While the methods used in the preceding pages suffice for most practical purposes, it is often very convenient to measure the intensities of association in different cases by means of some formula or 'coefficient'." We now know what Mr Yule considers "sufficient for most practical purposes"! Here are a few tables to illustrate it.

(i)

	A	Not-*A*	Totals
B ...	266,374	233,626	500,000
Not-*B* ...	233,626	266,374	500,000
Totals ...	500,000	500,000	1000,000

(ii)

	A	Not-*A*	Totals
B ...	934,579	31,153	965,732
Not-*B* ...	31,153	3,115	34,268
Totals ...	965,732	34,268	1000,000

(iii)

	A	Not-*A*	Totals
B ...	999,000	450	999,450
Not-*B* ...	450	100	550
Totals ...	999,450	550	1000,000

The values of Mr Yule's δ for these three tables are

(i) 16374, (ii) 1940, (iii) 100.

If anything is to be judged from these results, (i) is far more highly associated than (ii) and (ii) again than (iii). All are far more highly associated than the Datura, which we will call (iv), and then the order of association sinks from (i) to (iv) in a marked manner.

Now here are Mr Yule's coefficients of association put against his δ's

δ:	(i) 16374,	(ii) 1940,	(iii) 100,	(iv) 1·34,
Q:	(i) ·130,	(iv) ·282,	(ii) ·500,	(iii) ·996.

As the one series goes down the other goes up! What, we ask, can be learnt from Mr Yule on the subject of association, when his methods, "sufficient for most practical purposes," thus contradict themselves?

Here in (i), (ii) and (iii) we have kept the total frequency constant, but perhaps the most absurd side of Mr Yule's δ is manifest if we alter the total N of the observations. Suppose Bateson and Saunders had experimented with 8300 plants instead of 83, then δ would have been 134 instead of 1·34. The association is of course absolutely the same, but how would Mr Yule interpret his two δ's?

We regret having to draw attention to the manner in which Mr Yule has gone astray at every stage in his treatment of association, but criticism of his methods has been thrust on us not only by Mr Yule's recent attack, but also by the unthinking praise which has been bestowed on a text-book which at many points can only lead statistical students hopelessly astray.

(e) *Mr Yule's Assumption as to Absurdities which must arise if Normal Distribution be applied to the "Blind."*

Another interesting fallacy is developed by Mr Yule on p. 638 of his paper. He writes:

> "Consider for a moment what the assumption of normality of distribution would imply in any case where there is an increase of, say, the blind from one age-group to the next. This must imply either (1) a fall in the mean of the assumed variable character, goodness of sight, I suppose—if the standard deviation is constant or falling, or (2) an increase of the standard deviation if the mean is constant or rising. If the first occurs, then there must be some people in the later age-group who are *much more blind* than any people in the first, and fewer people of first-class sight; if the second, there must still be some people in the later group much blinder than any in the earlier, and there will also be some of much better sight. On the assumption that lies at the base of the normal coefficient, you cannot, in fact, effect a change in the numerical proportion of A's without changing them qualitatively at the same time. The assumption seems to me absurd, to be equivalent in this case to saying that there are certain people entirely deprived of sight in the first age-group, and certain others more than entirely deprived of sight in the second. The normal coefficient is accordingly inapplicable, and its precise values of no special significance."

We have rarely come across a more specious fallacy. If it were true it would be impossible in practical statistics to represent both a population and a selected sub-population by normal curves. Let the original population be N, mean M,

and standard deviation Σ, and let the selected population be n, mean m, and standard deviation σ. Then the second curve will always pass outside the first if

$$(m-M)^2+2(\Sigma^2-\sigma^2)\log_e\left(\frac{n\Sigma}{N\sigma}\right)>0.$$

This will happen of course if $\Sigma=\sigma$, Mr Yule's first case. In this case the curves cut on one side only of the means in the point

$$\breve{x}=\tfrac{1}{2}(m+M)+\frac{\log_e N-\log_e n}{m-M}\sigma^2.$$

In the case of blindness we should have $m>M$ *, and all it would signify would be that after this value, $\breve{x}$, of badness of sight, the older age would have for each grade of bad sight more individuals than at the lesser age; since the two curves *both extend to infinity* there is no question, as Mr Yule suggests, of persons being "much more blind" than at the younger ages. Mr Yule is simply confusing in his own mind or in the minds of his readers two senses of "more blind," i.e. more blind persons of each grade, and blind men of greater degree of blindness than actually can occur at the younger age. If he merely means to say that the Gaussian extends to infinity in both directions, that is a very old objection on the *theoretical* side to the curve; it has little value in practical statistics, where there is a reasonable approach to normality.

If $m=M$, then we have

$$\breve{x}=M\pm\sigma\sqrt{\frac{2\Sigma^2\{\log_e(\sigma N)-\log_e(n\Sigma)\}}{\sigma^2-\Sigma^2}}.$$

It is necessary therefore for real roots that $\sigma>\Sigma$, if σN be $>n\Sigma$. This is Mr Yule's second case, and this would give increasing numbers of persons in each grade of good sight beyond the value of $\breve{x}$ from M for increasing old age. We should have thought that this was a very improbable state of affairs, as it is almost certain that sight deteriorates with age after childhood at least. If we take the 1891 census data† used by Mr Yule, we have:

For ages: ...	*45—55*	*55—65*
Males in general ...	1,191,789	770,124
Blind males ...	1,752	1,905

If we suppose in these cases $m=M$, we find $X/\Sigma=2{\cdot}974$ and $x/\sigma=2{\cdot}810$, and since the dichotomy at "blind" is the same for both curves, we must have $X=x$,

* We have taken positive axis towards worse sight.

† Vol. III. pp. v. and lvii. Mr Yule has clubbed together those blind from childhood and the numbers, 12 times as great for these years, not blind from childhood. That so many *acquire* blindness indicates what a range of graduated sight there must be unless we suppose blindness to arise instantaneously.

or $\sigma = 1{\cdot}058\Sigma$. Thus a six per cent. increase in σ suffices to provide the increased blind population at the greater age; the two curves intersect at $x = \pm\, 2{\cdot}968\sigma$, or somewhat beyond the "blind" boundary. For all grades beyond this there will be more persons of each grade of bad sight. There is nothing inconceivable or improbable in this; but, on the other hand, there would be beyond $x = -\, 2{\cdot}968\sigma$, on the good sight side, a number of grades of better sight with more people at the older age. This, of course, is not impossible, but it seems far more reasonable to suppose the average sight to grow worse with old age, and in addition to change its variability somewhat. If the variability remained the same, as we have seen under the first case, there is no excess in the grades of marked good sight in the population of older ages. Let us now consider what happens if the mean be shifted and the variability increased. Taking $m > M$ *, we have

$$\frac{\breve{x} - m}{\sigma} = \pm \frac{\Sigma^2}{\sigma^2 - \Sigma^2}\sqrt{\frac{(m-M)^2}{\Sigma^2} + 2\frac{\sigma^2 - \Sigma^2}{\Sigma^2}\log_e \frac{\sigma N}{\Sigma n}} - \frac{\sigma\Sigma}{\sigma^2 - \Sigma^2}\cdot\frac{m - M}{\Sigma}.$$

Now let us take $\sigma = 1{\cdot}01\Sigma$, and therefore for the "blind" groups at 45–55 and 55–65 we have as before $X = 2{\cdot}9740\Sigma$, $x = 2{\cdot}8105\sigma = 2{\cdot}8386\Sigma$ to fix the dichotomy. Hence

$$(m - M)/\Sigma = (X - x)/\Sigma = {\cdot}13538.$$

Thus we find $\qquad (\breve{x} - m)/\sigma = +\, 2{\cdot}6738 \text{ or } -16{\cdot}2791.$

The former value shows that from some little distance outside the dichotomic line each grade of bad sight and blindness has more individuals of that grade at ages 55–65 than at ages 45–55. The latter value indicates that the point of intersection of the sight curves for the two ages on the side of good sight takes place at a point so extremely distant from the average sight that not a single individual would occur with such sight in a population many thousand times greater than the actual population.

We think it most probable, however, that a third case, not even referred to by Mr Yule, best describes what actually takes place—namely, that the sight at the older age gets worse and is *less* variable, not more variable. To illustrate this, take $\sigma = {\cdot}99\Sigma$, then $X = 2{\cdot}9740\Sigma$, $x = 2{\cdot}8105\sigma = 2{\cdot}7824\Sigma$. Hence

$$(m - M)/\Sigma = {\cdot}19158, \text{ and } (\breve{x} - m)/\sigma = +\, 2{\cdot}4742 \text{ or } +\, 16{\cdot}5875.$$

Thus the older age curve now never cuts the younger age curve of sight on the side of good sight at all. It cuts on the side of bad sight twice, once somewhat on the good sight side of the "blind" dichotomic line, and on the other occasion immensely beyond the limits of the populations in question. In other words, the older ages have fewer members in each grade of good sight and more members in each grade of very bad sight. This appears to us a perfectly reasonable state of affairs, and of course extends far beyond the ratio selected for σ/Σ.

It will thus, we think, be clear that had Mr Yule attempted to turn his half-baked notions into figures before he expressed them in words, he would

* The positive direction of the variate is towards bad sight.

have realised that more people in the more blind grades or more people in the better grades of sight are not the same thing as some people "much more blind" or some people of "much better sight" than any in the earlier aged groups. Mr Yule's alternatives are not real alternatives, he makes no reference to a shift in mean and a decrease in variability; such a combination involves only a reduction of the numbers in the grades of good sight, a very reasonable result with increasing age, and an increase of the numbers in the grades of bad sight, also a very reasonable hypothesis. Towards the 'tails' of both age curves, theoretically there would be fractional units, while in actual observations there would be isolated units at relatively wide intervals (cf. Galton's "Difference Problem *"). What the distribution of such units might be, could not be *a priori* predicted. But it is quite possible for two distributions with slightly separated means and slightly different variabilities to give quite reasonable fits to Gaussian curves and yet the distribution with the greater variability to have no outlying units with "much more" of a character than any which exist in the less variable distribution. The variabilities are much more closely determined by the bulk of cases with moderately large deviations than by the one or two extreme outlying individuals.

In the case we have last discussed the age-group 45–55 has two individuals lying outside $4{\cdot}65\Sigma$, and the distance between the means being $\cdot 19\Sigma$ this gives $4{\cdot}46\Sigma = 4{\cdot}51\sigma$ for the corresponding distance on the 55–65 age curve. Outside this limit are 2·5 individuals of this older age curve. Are we to say that that *half* individual represents Mr Yule's necessity for some people in the later age-group who are "*much more blind*" than any people in the first? In truth some people would be much less blind, if they would only stay to express their opinions in actual numbers before writing them down. The "minute sifting of numerical results" is the foundation of all true statistical inference, and here, as in other phases of his recent work, Mr Yule has committed himself to superficial statements reached by verbal disquisition which vanishes into nothingness if the touchstone of numerical investigation be applied to it.

There are many other points at which we should like to traverse Mr Yule's statements, but we think we have brought forward enough evidence to indicate how unreliable are his methods and how biased are his criticisms.

(16) *Summary of General Conclusions.*

In order to sum up the general conclusions reached in this paper, we must state first one or two principles which we accept as almost axiomatic:

(i) There is no universal method of dealing with an $n \times n$-fold table, except the method of mean square contingency, leading to a probability measure of the independence of the two characters, unless we know:

* *Biometrika*, Vol. I. p. 390.

(*a*) that both characters proceed by discrete units, and are tabled as such.

In this case the method of pseudo-ranks is identical with that of the product-moment, and there never has been doubt as to how the table is to be treated; or:

(*b*) that the frequency of the two characters is continuous and that this frequency follows or approximates to a definite theoretical system. There is only one such frequency system, which has up to the present been effectively discussed, i.e. that of Laplace, or as it is more frequently but less justly called, that of Gauss*. If the distribution be Gaussian or approximately Gaussian, there are many ways of dealing with an $n \times n$-fold table.

(ii) A majority of the cases which occur in statistical practice are so close to the Gaussian distribution that methods based upon Gaussian theory will give useful first approximations, i.e. correlations within $\pm$ ·05 say of the true values.

Years ago one of the present writers insisted on the non-Gaussian character of many variables. But he also remarked on the large number of variables which can be described with sufficient practical accuracy by a Gaussian distribution.

The present discussion demonstrates that even with distributions markedly skew the Gaussian theory, if applied to 2×2-fold tables—without extreme dichotomies—will give results not differing by more than ·05 from the value of the true correlation and often differing by much less. Roughly, we may say that for reasonable divisions, the divergence between the true correlation and that obtained by Gaussian theory is hardly ever of practical importance and indeed in "populations" of the size usually dealt with rarely exceeds twice the probable error.

(iii) The coefficient of correlation has such valuable and definite physical meanings that if it can be obtained for any material, even approximately, it is worth immensely more than any arbitrary coefficients of "association" and "colligation."

Starting from these principles we ask ourselves to what data Mr Yule proposes to apply his three processes;

(*a*) The Boas-Yulean ϕ for fourfold tables.

(*b*) The coefficient of pseudo-ranks.

(*c*) The coefficient of association or that of colligation.

We have shown in this paper that for tables with a finite number of cells of the order 5×5-fold to 8×8-fold, the method of pseudo-ranks must lead to a value below and often 40 °/ₒ below the true correlation of variates. Mr Yule has stumbled into a statistical pitfall, for he has neglected the fact that correlation of ranks is not correlation of variates, and that his correlation of ranks would still have to be corrected for the class-index correlations, i.e. he has also neglected the

* Of course both these writers only dealt with the frequency of one variate; Bravais extended it to two, but gives no admissible proof of his formula, which he practically gets by analogy.

existence of huge brackets. We may, we hold, entirely dismiss from statistical practice this method of pseudo-ranks, except for the case wherein it has always been used, i.e. for discrete unit variates, classified by units, where it coincides with the usual product-moment method and needs no special name.

We have only to note here that Mr Yule uses the method of pseudo-ranks, which we hold to be demonstrably false, to make very sweeping charges, which can be and have been met, against the pigmentation work of the Biometric School. He not only suggests that the workers on the pigmentation data were foolish, but that they were dishonest. That is the sort of attack which usually recoils on the head of the man who makes it, especially when he has for several years worked in the Department against which he prefers the accusation. As a matter of fact the non-Gaussian character, the variability of tetrachoric r_t, for different divisions was recognised very soon after it had been applied. But the investigations then made and more amply illustrated in this memoir indicate that the values originally given were substantially correct, the inheritance of intensity of pigmentation between parent and offspring lies between ·46 and ·50; it is not of the order $\frac{1}{3}$ as Mr Yule asserts on the basis of a theory which we feel convinced he will have to withdraw, if he wishes to maintain any reputation as a statistician.

We have shown in the course of this memoir that the coefficient of association Q, if treated merely as an undefined measure of association, has not for varying dichotomies the stability of the tetrachoric coefficient and it appears to have no reasonable physical meaning even for the cases which he has selected. Mr Yule has deduced from it a second coefficient, that of colligation, which has, he says, a physical meaning, when the table has been dressed in an artificial manner, namely it signifies on this artificial table the excess of the percentage of A's that are also B's over not-A's that are also B's. We have shown from Mr Yule's own writings that such a difference of percentages has in his own practice no meaning at all from the standpoint of association.

Mr Yule never tells us clearly when we are to use one or other of his coefficients. He spends 16 pages of his memoir on discussing the application of his coefficients of colligation and association to the vaccination data; yet on p. 611 he writes: "For discontinuous attributes—attributes proper, as we might term them—the true correlation is that given by formula (24) or (26) [i.e. the Boas-Yulean coefficient or Pearson's ϕ]; we are dealing with a variable in fact, which can only take two values as distinct from a variable exhibiting a normal or any other continuous distribution. Tables I, III and IV [i.e. the vaccination data], as it seems to me, represent precisely such a case." Here Mr Yule has given up colligation as applied to vaccination; if so why devote 16 pages to its discussion? But 20 pages later Mr Yule tells us that: "For investigations on smallpox and vaccination such as those of Brownlee and Macdonell and Turner, the use of Q or ω would, in my opinion, have been more illuminating as well as simpler than the use of the normal coefficient" (p. 631).

Mr Yule does not really appear to know which of his ducklings to prefer, even for what in his estimation—although not in ours—are discrete attributes. We are quite clear that none of them are appropriate. Still less are they appropriate for the definitely continuous data to which Mr Yule and his disciples apply them.

The controversy between us is much more important than an idle reader will at once comprehend. It is the old controversy of nominalism against realism. Mr Yule is juggling with class-names as if they represented real entities, and his statistics are only a form of symbolic logic. No knowledge of a practical kind ever came out of these logical theories. As exercises for students of logic they may be of educational value, but great harm will arise to modern statistical practice, if Mr Yule's methods of treating all individuals under a class-index as identities become widespread, and there is grave danger of such a result, for his path is easy to follow and most men shirk the arduous.

The very large amount of arithmetic involved in this paper would have been impossible without friendly help from a number of our colleagues; we have especially to thank Miss Julia Bell and Mr Herbert E. Soper; the former for much work in calculating and the latter for diagram draughtsmanship as well as calculation. Miss Ethel M. Elderton also most kindly undertook one or two pieces of heavy arithmetic. We can hardly hope to have escaped numerical slips, but we feel confident that such slips, if they occur, will as frequently tell against us as for us; and we have not, knowing the fallibility of the best calculators, done more than draw attention to points where our arithmetic differs from that of Mr Yule. We have laid sole stress on errors of interpretation and on fallacious theory.

APPENDIX I.

On the Fallacy of asserting Perfect Association when one Quadrant in a Fourfold Table is Vacant.

In all the values of the probable errors hitherto determined the constants of the distribution found in the formula are truly constants of the *actual* distribution of which the population under discussion is a *sample*. Because we do not know the actual distribution, we replace its constants by those of the observed sample. This method will as a rule not lead us astray, but it may do so grievously in cases where the observed frequencies take limiting values. For example, consider the population described in the following fourfold table:

TABLE I.

	A	Not-A	Totals
B... ...	971,138	22,862	994,000
Not-B...	5,862	138	6,000
Totals	977,000	23,000	1000,000

and let this distribution be the real population or an absolutely proportional sample. Now for this population with the scheme $\frac{a \mid b}{c \mid d}$, we have $ad - bc = 0$, and all our measures of association vanish. If we take a sample of n' from such a population of n, the value $\eta = (ad - bc)/n^2$ will not *for the sample* be zero, but, if

$$\eta' = (a'd' - b'c')/n'^2,$$

where a', d', b' and c' are the values in the sample, will have a standard deviation*

$$\sigma_\eta = \sqrt{\frac{(a+b)(a+c)(d+b)(d+c)}{n^4}} \times \frac{1}{\sqrt{n'}}.$$

Now let us compare this with

$$\sigma_{\eta'} = \sqrt{\frac{(a'+b')(a'+c')(d'+b')(d'+c')}{n'^4}} \times \frac{1}{\sqrt{n'}},$$

the value of $\sigma_{\eta'}$ derived from the sample itself. Now to *units* the value of the sample might be

TABLE II.

	A	Not-A	Totals
B ...	971	23	994
Not-B ...	6	0	6
Totals	977	23	1000

and it is clear that σ_η and $\sigma_{\eta'}$ will be exactly the same, as they depend only on the marginal totals. Thus

$$\sigma_\eta = \sigma_{\eta'} = \frac{1}{\sqrt{1000}} \sqrt{\cdot 994 \times \cdot 977 \times \cdot 006 \times \cdot 023}$$
$$= \cdot 000366.$$

But $\eta' = \cdot 000138$. Accordingly $(ad - bc)/N^2$ when $d = 0$ is not significant, having regard to its probable error, and the association is zero. On Mr Yule's theory $Q = -1$ and its probable error is zero.

Now the standard deviation of Q, for Q zero, is

$$\frac{1}{\sqrt{n'}} \sqrt{\frac{(a+c)(c+d)(b+d)(a+b)}{(ad+bc)^2}},$$

for a sample of n', where a, b, c, d refer to the original population.

Let us consider the differences for the above material which will arise from calculating ${}_0\sigma_Q$ on the population in Table I and on the sample in Table II. On Table I we find ${}_0\sigma_Q = 1\cdot 37$; on Table II ${}_0\sigma_Q' = 2\cdot 62$. Hence either Table is

* Pearson, "On a novel method of Regarding the Association of two Variates," *Drapers' Research Memoirs, Biometric Series*, No. VIII, Dulau & Co., 1912, p. 7.

absolutely compatible with Q *being zero*. It is true that the value of ${}_0\sigma_Q$ as estimated from the sample is nearly double what its value would be truly estimated from the population, but both values suffice to show that $Q=0$ is as reasonable an hypothesis as to the constitution of the material as $Q=-1$, and far more reasonable than $Q=-1\pm 0$. The method of mean square contingency gives us $\chi^2=\cdot 1455$, leading by Palin Elderton's Tables to $P=\cdot 971$, or if the material were truly independent only in three cases per hundred should we get a better fit than Table II to Table I in taking samples of 1000.

The fact is that in drawing random samples of 1000 from Table I, the distribution of frequency in the cell d is given by

$$(\cdot 999862+\cdot 000138)^{1000}=\cdot 8712+\cdot 1202+\cdot 0086 \text{ (for all terms beyond second).}$$

Hence in 100 samples of Table I, d would be zero in 87 cases, unity in 12 and greater than unity in about one case. In other words Mr Yule's Association Coefficient would for material with true zero association be -1 in about 87 °/₀ of cases and of the order $+\cdot 75$ in 12 °/₀ other cases. In all these cases, however, the probability method shows practical independence. It must accordingly be recognised that it is extremely dangerous, if zero frequency be found in one quadrant, to assert that $Q=\pm 1\pm 0$. For, a population of zero association would give such a value in 87 °/₀ of samples of 1000 in a case like that under consideration. The tetrachoric method fails also for this extreme case, for the simple reason that the very continuity of the method excludes thinking in isolated units. If the real population were that of Table I, then on the basis of continuity we should expect $\cdot 138$ from the infinite skirt of the Gaussian surface in quadrant d. But we might extend this volume of the skirt up to something under $\cdot 5$ before we should anticipate a whole unit to appear in d. There is no suggestion of this kind about Q, for Mr Yule directly discards all conception of such continuous frequency surfaces. This point must be borne in mind in applying tetrachoric r_t; a quadrant of zero frequency does not necessarily signify that in the theoretical Gaussian distribution this quadrant would have zero frequency. We may equally reasonably assume it anything up to $\cdot 5$. In such a distribution as

TABLE III.

	A	Not-A	Totals
B ...	21	450	471
Not-B ...	529	0	529
Totals	550	450	1000

we shall alter the tetrachoric coefficient from negative unity to a high value somewhat less than negative unity, by inserting anything up to $\cdot 5$ in quadrant d, but we shall not swing the correlation from -1 through 0 to a small positive value by the process. We shall get a reasonable minimum value for r_t, and we shall be convinced that $r_t=-1\pm 0$ is not necessarily a true representation of the

state of affairs, although the correlation is negatively high. But in such a table as Table II the whole character of the correlation is changed by inserting a frequency x, less than ·5, in this d quadrant.

The first question to be answered is: If we put a small frequency x in d, how are we to take it from the other quadrants? The only reasonable answer here seems to be: "Take it in proportion to their frequencies." Thus Table II becomes of the form:

TABLE IV.

$971\left(1-\frac{x}{1000}\right)$	$23\left(1-\frac{x}{1000}\right)$	$994\left(1-\frac{x}{1000}\right)$
$6\left(1-\frac{x}{1000}\right)$	x	$x+6\left(1-\frac{x}{1000}\right)$
$977\left(1-\frac{x}{1000}\right)$	$x+23\left(1-\frac{x}{1000}\right)$	1000

Now can we choose x so that $ab-cd=0$, for this Table? For this we must have:

$$971\left(1-\frac{x}{1000}\right)\times x=23\left(1-\frac{x}{1000}\right)\times 6\left(1-\frac{x}{1000}\right),$$

which leads to $x=·1421$.

This shows us that on the Gaussian hypothesis a value of x absolutely consistent with a zero being recorded in that quadrant leads to zero association. This for most practical purposes would be sufficient. But a little further consideration here is desirable; we may write the table in the form:

TABLE V.

$970·8620-x$	$22·9967+x$	993·8587
$5·9992+x$	$·1421-x$	6·1413
976·8612	23·1388	1000

Here $x=0$ gives a table of zero association, $x=·1421$ gives, if we can only record to a unit individual, the table of experience, i.e.

TABLE VI.

971	23	994
6	0	6
977	23	1000

Now η of Table V $=-\dfrac{x}{1000}$. Therefore the probable error of $x=1000$ times that of η. If we find ${}_0\sigma_\eta$ for $x=0$ in Table V, we have for its value ·0003714,

quite close to the value we obtained for Table I from which we originally deduced Table II. Hence the probable error of $x = \cdot67449 \times 1000 \times {}_0\sigma_\eta = \cdot2505$. It will be clear therefore that with such a probable error for x, the frequency of the d quadrant might actually be anything from 0 to ·5, and yet the Table of experience takes the form VI, when we can record only by units.

Suppose d had been ·5, $x = -\cdot3579$, then we have for the theoretical table:

TABLE VII.

971·2199	22·6388	993·8587
5·6413	·5000	6·1413
976·8612	23·1388	1000

which solved by the tetrachoric process gives

$$r_t = + \cdot224 \pm \cdot187,$$

i.e. the value of r_t is not significant*.

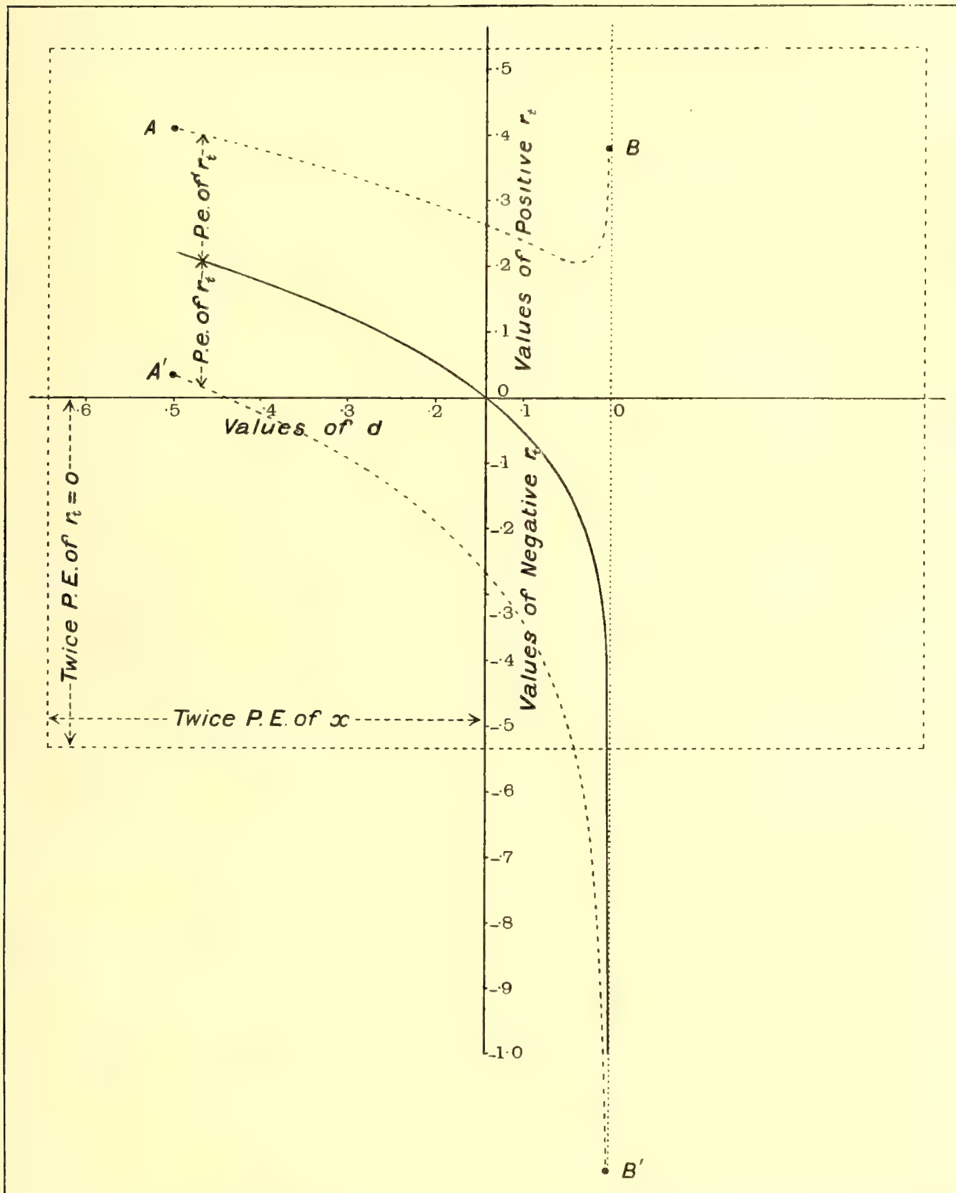
On the other hand if we put $d = 0$, i.e. deal with Table VI, tetrachoric r_t is unity, because on the Gaussian hypothesis only complete association is compatible with absolute zero in this quadrant and then any sample will exhibit absolute correlation. But, as we have just seen, the zero in quadrant d could arise from samples of a population in which d was small but not zero, and that in a particular case 87 °/₀ of samples of a material with zero association would show this zero quadrant. The accompanying diagram gives the values of tetrachoric r_t for various hypotheses as to x. The dotted rectangle is bounded by the lines which give vertically twice the probable error of x for $x = 0$, and horizontally twice the probable error of r_t for true $r_t = 0$. We have also placed *once* the probable error of tetrachoric r_t from the plotted r_t on either side giving the broken curve. It will be seen that from the case of $d = \cdot5$, $r_t = + \cdot224 \pm \cdot187$ up to $d = \cdot0025$ and $r_t = - \cdot400 \pm \cdot779$, there is no significance in the values of r_t found by the tetrachoric process. Even after this we cannot assert that the values of r_t obtained would be significant†; for the proof of the formula of the probable error of r_t depends upon δd being small as compared to d. Now $\sigma_d = \sqrt{d(1 - d/N)} = \sqrt{d}$, very nearly when d is very small. Hence $\delta d/d$ is of the order $\sqrt{d}/d = 1/\sqrt{d}$ which will be large if d be less than unity‡. The probable error of $r_t = 1$ is only zero if d is absolutely zero for the population which is being sampled and not if it is merely zero in the sample. But we only know that population through the sample, and we see that in such cases as we are considering the zero in the sampled population is only likely to occur in a small proportion of the cases dealt with. The tetrachoric process clearly fails in such cases, but we see that with

* If we apply Q we have $Q = \cdot584 \pm \cdot332$, pointing rather more in the direction of significance, but such application of Q is illegitimate, as the fractionising of the *theoretical* surface has no meaning for a coefficient which is based on complete neglect of the nature of the frequency-distribution.

† In fact the tetrachoric r_t series-equation rapidly becomes divergent and the formula for the probable error takes an indeterminate form.

‡ The like failure occurs in Mr Yule's proof for the probable error of Q, although he has not warned his readers of this; it is accordingly not applicable, if one quadrant has zero frequency.

any reasonable assumption with regard to x, the tetrachoric process suggests not perfect but zero association; the perfect association is only reached as a limiting case, although Mr Yule's coefficient gives an unhesitating unity*.



* In the Table selected by Mr Yule to exhibit the variations of his coefficient of association (*Phil. Trans.* Vol. 194 A, p. 277) 51 out of the 164 values shown are ±1 owing to the occurrence of a zero quadrant. With the introduction of a ·5 into this quadrant—quite reasonable on the basis of a continuous variate—his values may be reduced from +1 to +·72 or swing over from −1 to +·98, etc., etc.! What is the value in such cases of the statement that there is "perfect association"?

We have seen that if the sampled population be really one of zero association, the sample may have in quadrant d in 87 °/₀ of cases zero frequency, but in 12 °/₀ of cases unit frequency. This unit must—working to unit individuals—be withdrawn from a, b and c. The most probable case is the addition of a unit to both a and d and their withdrawal from b and c, or again we may draw units respectively from a, b or c. We may consider only these four cases:

(α)

972	22	994
5	1	6
977	23	1000

(β)

971	22	993
6	1	7
977	23	1000

(γ)

971	23	994
5	1	6
976	24	1000

(δ)

970	23	993
6	1	7
976	24	1000

These lead to

	α	β	γ	δ
Value of tetrachoric r_t	$\cdot 39 \pm \cdot 15$	$\cdot 36 \pm \cdot 16$	$\cdot 38 \pm \cdot 15$	$\cdot 35 \pm \cdot 16$
P. E. of $r_t = 0$ for same marginal totals...	$\cdot 00 \pm \cdot 27$	$\cdot 00 \pm \cdot 25$	$\cdot 00 \pm \cdot 27$	$\cdot 00 \pm \cdot 25$
Value of Association Q	$\cdot 80 \pm \cdot 14$	$\cdot 76 \pm \cdot 16$	$\cdot 78 \pm \cdot 13$	$\cdot 75 \pm \cdot 16$
P. E. of $Q = 0$ for same marginal totals	$\cdot 00 \pm \cdot 23$	$\cdot 00 \pm \cdot 24$	$\cdot 00 \pm \cdot 23$	$\cdot 00 \pm \cdot 25$

It will be seen that r_t is always less than 2·5 and Q always greater than 4·5 times the probable error. The value of r_t obtained is always less than 1·5 times the probable error of $r_t = 0$ for the same marginal totals, while Q is 3 times the probable error of $Q = 0$ for the same marginal totals. Thus the tetrachoric method warns us that the association is probably zero, while the association coefficient emphasises a high value of the association.

To sum up: The correct process in these cases with zero in one quadrant is not to assert that the association is perfect with Mr Yule, but to apply first the probability test and determine whether the material may not rather be a random sample from an original population of zero association. The failure of the tetrachoric r_t in these cases is rendered evident in the working, we reach non-convergent series and are thrown back on a limiting case; if we place in the zero quadrant a small frequency less than 0·5, which would correspond to zero in the actual table, we find a finite value of r_t, but one non-significant having reference to its probable error, unless we approach close to the limit of $d = 0$, in which case the probable error of r_t is so far undetermined, because the ordinary process fails to be valid.

Mr Yule's view that association is perfect when there is zero in any quadrant ignores the fact that he can only deal with a *sample* of the true population, and

so far as we are aware he has not warned students of the extreme danger of his Q in these cases. The non-applicability of tetrachoric r_t to cases in which one of the quadrants has zero or small relative frequency is, on the above and on other grounds, well-known to those who habitually work with it, and the rule has long been to avoid where possible all such extreme dichotomies. This point has been largely disregarded by Mr Yule in his criticism of tetrachoric methods.

APPENDIX II.

ON THE TEST OF GOODNESS OF FIT OF OBSERVATION TO THEORY IN MENDELIAN EXPERIMENTS.

An objection to the χ^2, P, test of goodness of fit, has recently been raised with somewhat unconscious humour by certain ardent Mendelians. A simple illustration may be taken. Suppose the mating to be

$$(DR) \times (DD) = 50\ {}^{\circ}/_{\circ}\ (DR) + 50\ {}^{\circ}/_{\circ}\ (DD).$$

For example, let there be 1000 offspring and let 480 of these be recognised by later experiment as (DR)'s and 520 as (DD)'s. Then the standard deviation is

$$\sqrt{1000 \times \tfrac{1}{2} \times \tfrac{1}{2}} = 15 \cdot 8$$

and the observed deviation is 20, and P = about ·90, or the fit is quite good. But suppose the observations are

		480 (DR)'s	519 (DD)'s and	1 (RR)
i.e.	Observation :	480	519	1
	Theory :	500	500	0

Clearly $$\chi^2 = S\frac{(\text{observation} - \text{theory})^2}{\text{theory}} = \infty,$$

and $P = 0$, or the probability of observation being a random deviation from theory is zero, there is absolute badness of fit. Hence either theory or observation is at fault. It is not, however, the theory of "goodness of fit" which fails in such a case, but the Mendelian theory which wants mending. If we put a black balls, b white balls and c red balls in a bag, the theory will tell us whether any sample of black, white and red balls can be reasonably considered as a random extract from that bag. But if we are presented with a series of black, white, red, and *green* balls and asked what is the probability that these were drawn from that bag, we must assert that it is zero, however concordant with theory the results for the black, white and red balls alone may be. There are only two courses open to the

experimenter who gets green balls however few in number out of a black, white and red ball theory:

(i) to assert that they are anomalies due to his temporary colour-blindness, to mistakes in his observations, or to some misinterpretation of his results, which are not accurately describable in terms of his original categories, or

(ii) to amend his original theory by inserting some green balls *a posteriori* into the bag and starting afresh to calculate his probabilities.

The one course weakens the weight we must lay on his record or on his choice of categories; the other tends to discredit his *a priori* theory. To adopt a third course and assert that we want a new test for "goodness of fit" of theory to observation, which shall cover such discrepancies, i.e. which shall slur over divergencies between *a priori* theory and *a posteriori* results, may appeal to our sense of human fallibility but scarcely to our appreciation of scientific logic. We shall be left with the suspicion that the theory is plastic and the observations elastic. What criterion of "goodness of fit" can the theory of probability provide when it is a case of applying plastic theory to elastic observations? The answer surely is none whatever until the plasticity of the theory has been quantitatively studied, and until the errors of the record have been quantitatively stated. Either we must be told that the observer will mistake a red ball for a green one in so many per cent. of cases, or we must be told that the theory will be inaccurate in so many per cent. of cases. Personally we think it possible that all attempts to find a "good fit" of a plastic theory to elastic observations are idle. It is a consideration of the green balls, which are said not to be in the black, red and white bag at all, which is often the basis of marked scientific progress. That atmospheric nitrogen differed from pure nitrogen was just such a "green" ball; but the plastic theory that air consisted of oxygen and nitrogen *only* had been confirmed by many elastic observations before Lord Rayleigh followed up his "green" ball.

These remarks are suggested by the following paragraphs in two recent Mendelian publications. Dr Raymond Pearl writes* concerning Mendelian data:

"A determination might be made of the 'goodness of fit' of theory to observation by Pearson's method, were it not for the fact that that method cannot be applied to cases like the present."

Dr Pearl says it cannot be applied because he finds *green* balls, where his theory puts only black, white and red into his bag. It is either his observation record or his Mendelian theory, not the mathematics of "goodness of fit," which needs modification. Dr Pearl continues in a footnote as follows:

"The difficulty lies in the fact that Pearson's test depends upon a variable

$$\chi^2 = S\left\{\frac{(m_r - m_r')^2}{m_r}\right\},$$

* *The Journal of Experimental Zoology*, Vol. XIII. p. 203.

where m_r is the theoretical frequency and m_r' the observed. Now obviously in any distribution where even one m_r is zero, the value of χ^2 must be infinite whatever may be the values of the other m_r's or m_r''s. That is, if the theoretically expected frequency on any base element is numerically zero, the probability against the whole curve becomes infinite. Thus, for example, suppose a system of frequencies like the following, a type which is continually arising in Mendelian work:

Class	*1*	*2*	*3*	*4*	*5*
Theoretically expected frequency ...	595	827	68	0	96
Actually observed frequency ...	594	828	67	1	96

"Now it does not need a mathematical measure of any kind to tell one that in this case the theoretical and actual distributions are in very close agreement. Yet because the *theoretical* frequency on class 4 is zero, the probability by Pearson's test is literally infinite *against* the observed distribution being regarded as a random sample of a population distributed in accordance with the theoretical frequencies. Pearson has indeed himself noted what is essentially this same difficulty in using the test on ordinary frequency distributions*."

Now what does this paragraph exactly signify? Interpret it in coloured balls; white, red, black and yellow balls are placed in a bag in large numbers in the proportions of 595 : 827 : 68 : 96. There are *no* green balls in the bag. One such green ball is said to have been drawn. Theory says it is an impossibility, and the criterion of goodness of fit says its improbability of occurrence is infinite. We can conceive no logical theory doing anything else. Dr Pearl does in fact tell us that cases in which a ball, classified as green, comes out of the *theoretical* white, red, black and yellow ball bag are of "a type which is continually arising in Mendelian work." This at any rate is a frank admission. As a matter of fact with his arbitrary division between "over 30-egg" and "under 30-egg" hens, we are not surprised that "green" balls appeared not only in ones, but in twos and even in fours, and in a few cases to even more, although this is attributed to "physiologically extremely favourable" matings (*loc. cit.* p. 248) as apart from gametic theory†. What Dr Pearl is seeking is a plastic theory or an elastic record, not a real criterion of goodness of fit, which must give no finite probability when green balls come from a bag which contains no green balls! Dr Pearl continues:

"The point noted obviously limits greatly the applicability of Pearson's test, and in a most unfortunate direction. Tests of goodness of fit are much needed in Mendelian work [we cordially agree!]. But it is just here that the classes where

* This is a complete misunderstanding. Pearson says that you must not in the case of *continuous* variation make use of classes which *theoretically* have each less than *unit* frequency, where the record goes by individual units only.

† Until Dr Pearl publishes the actual record of each bird, and not merely its class-index, and the same for its ancestors, it is impossible to estimate his degree of justification for the theoretical treatment of his results.

the theoretical frequency is zero often occur*." In other words, observation records something as occurring which existing theory says cannot occur, and Dr Pearl asks for a criterion which shall make the impossible only mildly improbable. He must either remould his theory or explain away his observations. We see no alternative.

Dr Pearl's demand for a criterion which shall not be crucial, but allow elastic records to fit a plastic theory, is well illustrated by the following paragraph from the work of another American Mendelian who is convinced that "feeble-mindedness" is a Mendelian unit character, but has found his "green ball" in the normal offspring of two feeble-minded parents:

"These two are apparent exceptions to the law that two feeble-minded parents do not have anything but feeble-minded children†. We may account for these two exceptions in one of several ways. Either there is a mistake in calling them normal, or a mistake in calling the parents feeble-minded; or else there was illegitimacy somewhere and these two children did not have the same father as the others of the family. *Or we may turn to the Mendelian law and we discover*

* The remainder of Dr Pearl's paragraph runs: "To determine the probable error of the individual frequency in measuring the goodness of fit of Mendelian observation and theory, as was first practised by Weldon, and later by Johannsen and by Mendelian workers generally, does not appear to the writer to be an altogether sound procedure. It fails to take account of the *correlations* in errors amongst the several frequencies. Yet these are just as important and just as certainly existent in a Mendelian 'category' type of distribution as in the ordinary variation polygon of a continuously variable character....Pearson's test covers this point, and were it not for the other difficulty noted above would be much more widely useful in Mendelian work than is actually the case" (*loc. cit.* p. 204). The "widely useful" test in Mendelian work is quite obviously one which will overlook negation of theory, or not drive the observer back to question the validity of his records or his categories. But there is a misstatement in the above sentence which needs correction. If there be only alternative categories, e.g. the total of (RR)'s compared with the total of (DR)'s$+(DD)$'s, then Pearson's test is absolutely identical with the probable-error test. This is of course well recognised; for, if m_1 and m_2 be the observed frequencies and n_1 and n_2 the theoretical frequencies,

$$\chi^2=\frac{(m_1-n_1)^2}{n_1}+\frac{(m_2-n_2)^2}{n_2}=(m_1-n_1)^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$=\frac{(m_1-n_1)^2}{n_1\left(1-\frac{n_1}{N}\right)}=\frac{(m_2-n_2)^2}{n_2\left(1-\frac{n_2}{N}\right)},\ \text{where } N=n_1+n_2,$$

$$=\frac{(\text{Deviation of either category})^2}{(\text{Standard Deviation})^2},$$

and for this case $P=\frac{2}{\sqrt{2\pi}}\int_\chi^\infty e^{-\frac{1}{2}\chi^2}\,d\chi$. This test therefore Weldon applied with perfect legitimacy to the consideration of Mendelian quarters. When Weldon came *in the very paper cited by Dr Pearl* to test more complex Mendelian results, he did not fail to take account of correlations in errors, and actually applied Pearson's criterion (*Biometrika*, Vol. I. p. 235). Dr Pearl's sentence therefore requires remodelling; he has clearly failed to appreciate what Weldon was doing.

† The pedigrees published by Weekes and Goddard show other exceptions to Mendelian law, which they appear to have overlooked. This is confirmed by similar pedigrees in our Laboratory. The very idea that the continuous and highly variable character "feeble-mindedness" is a "unit character" in the Mendelian sense will do much to check real research into the grave complexities and difficulties of this very vague and broad category.

that according to that law there might be in rare instances such a combination of circumstances that a normal child might be born from two parents that function as feeble-minded. For practical purposes it is, of course, pretty clear that it is safe to assume that two feeble-minded parents will never have anything but feeble-minded children*."

The italics are ours. They are very typical of the manner in which an elastic record and a plastic theory are made to fit. No account is anywhere provided of this extension of Mendelian theory which "in rare instances" provides a normal offspring to the two parents with an abnormal dominant character. Possibly, as in Dr Pearl's case, it is due to the occurrence of "physiologically extremely favourable" matings. Anyhow the old definite simplicity of Mendel's Mendelism has gone; with an elastic record and a plastic theory any data may be Mendelian —or not—according to the views of the investigator who moulds his theory and stretches his facts.

How welcome to such an one must be the Yulean theory of association! "Whatever the nature of the classification, however, natural or artificial, definite or uncertain, the final judgment must be decisive; any one object or individual must be held either to possess the given attribute or not" (Yule, *Theory of Statistics*, p. 9).

In the face of such a direction, how could Dr Pearl have been so foolish as to balance a number of his hens on the dichotomic fence of 30-eggs, and allow a moiety of each such hen to possess one Mendelian attribute and the other moiety its alternative?

It is not difficult to understand, however, why Dr Pearl does not like Pearson's criterion of the goodness of fit of theory and observation. On p. 255 of his memoir he gives a table "showing the observed and expected distributions of winter egg production for all matings taken together." He remarks on this table that "the lumped figures do not give an altogether fair estimate of the matter, but some sort of a summary is necessary." We agree very cordially because a number of the impossible green balls of the subsidiary tables do not appear as such when the tables are lumped, but taking the data for what they are worth we have the following four series:

		Winter Egg Production		
		Over *30*	Under *30*	Zero
I.	Observation	365·5	259·5	31
	Theory	381·45	257·25	17·30
II.	Observation	2	23	15
	Theory	0	25	15
III.	Observation	36	79	8
	Theory	26·5	86·75	9·75
IV.	Observation	57·5	98·5	23
	Theory	68·60	95·0	15·40

* *The Kallikak Family*, p. 114.

Pearson's criterion gives the following results:

I. $P = \cdot 003$
II. $P = \cdot 000$
III. $P = \cdot 113$
IV. $P = \cdot 036$

The odds against the first series arising from material following the theory are 332 to 1; the second series is impossible on the theory; the odds against the third series are about 8 to 1; and against the fourth series are about 27 to 1. The combined odds against even the three series (I, III and IV) representing the theory are very large indeed. Dr Pearl actually tells us that "the investigator is usually expected to reject abnormal material" (p. 256). And he prides himself on not having done so*, and asks us to form a judgment not on the summary but on the detailed data in the body of the paper. We have done so, and the criterion gives still worse results. We agree with Dr Pearl that "the high producing hen, somewhat like the race horse, is a rather finely strung, delicate mechanism, which can be easily upset, and prevented from giving full normal expression to its inherited capacity in respect to fecundity" (*loc. cit.* p. 255). But surely this is only to admit that the character chosen was wholly unfit to test the theory upon at all? It does not justify rejecting the only scientific test of "goodness of fit," and then concluding from nothing other than general impression that "the cumulative probability that the hypothesis applied represents at least a reasonable approximation to the true interpretation of the results becomes very great" (*loc. cit.* p. 257). If "cumulative probability" signifies anything at all, it means the theory of probability applied to the series to deduce combined odds against the total results and these are hopelessly against Dr Pearl. Further we cannot go until Dr Pearl publishes his record, which is not yet before us, although he has published his own interpretation of it in a great variety of journals.

APPENDIX III.

On the Equation to the Surface of Constant Q.

If the limits to the frequency range in the variate x be a and a', and in the variate y be b and b', then in the notation of p. 184

$$n_{pq} = \int_x^a \int_y^b z\,dx\,dy,$$

$$p = \int_x^a \int_{b'}^b z\,dx\,dy, \quad q = \int_{a'}^a \int_y^b z\,dx\,dy,$$

$$z = \frac{d^2 n_{pq}}{dx\,dy}.$$

* "But in view of the rather hysterical attacks upon geneticists and their method of work in this country, if for no other reasons, it seems best to follow the plan of publishing all the data." We would remind Dr Pearl that this is exactly what he has *not* done. We require the quantitative record of every individual hen and its ancestry as far as is known before we can fully test the validity of his results.

By the equation in the second line of p. 184, we have

$$n_{pq} = \frac{(1-\chi)(p+q) + N \pm \sqrt{\{N-(1-\chi)(p+q)\}^2 + 4\chi(1-\chi)pq}}{\frac{1}{2}(1-\chi)}.$$

Whence by straightforward but somewhat laborious differentiation we find

$$z = \frac{\dfrac{dp}{dx}\dfrac{dq}{dy}\chi\{N^2 - N(1-\chi)(p+q) + 2(1-\chi)pq\}}{\{[N-(1-\chi)(p+q)]^2 + 4\chi(1-\chi)pq\}^{\frac{3}{2}}} \quad \text{............(i).}$$

This is the equation to the surface of constant Q. As soon as we know the nature of the marginal frequencies, i.e. the values of p and q, we can find the form of the surface. The above equation is somewhat simplified if we refer p and q to the medians of their frequency-distributions, i.e. write $p = \frac{1}{2}N - \alpha$, $q = \frac{1}{2}N - \beta$. In this case

$$z = \frac{\dfrac{d\alpha}{dx}\dfrac{d\beta}{dy}\chi\{\frac{1}{2}N^2(1+\chi) - 2(\chi-1)\alpha\beta\}}{\{\chi N^2 - (\chi-1)[(\alpha+\beta)^2 - \chi(\alpha-\beta)^2]\}^{\frac{3}{2}}} \quad \text{... ....(ii).}$$

If the marginal frequencies are Gaussian,

$$\alpha = \frac{N}{\sqrt{2\pi}\sigma_1}\int_0^x e^{-\frac{1}{2}\frac{x^2}{\sigma_1^2}}dx, \quad \beta = \frac{N}{\sqrt{2\pi}\sigma_2}\int_0^y e^{-\frac{1}{2}\frac{y^2}{\sigma_2^2}}dy,$$

$$\frac{d\alpha}{dx} = \frac{N}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\frac{x^2}{\sigma_1^2}}, \quad \frac{d\beta}{dy} = \frac{N}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\frac{y^2}{\sigma_2^2}}.$$

It is therefore possible by aid of Sheppard's Tables to construct the contour lines of the surface of constant Q for this relatively simple case. But the surface is far from simple and its complex equation seems to indicate that association as measured by Q is of a very arbitrary character. We have constructed the surface of constant association for the special case of Gaussian marginal frequencies, when $Q = \cdot 6$. The photograph of the surface, the regression lines and the contours will be published on another occasion. It suffices here to note: (i) that the arrays are heteroscedastic, varying from homoscedasticity of the mid-section to a skewness of ·16 when $x/\sigma_1 = 1 \cdot 5$ and to a skewness of ·20 when $x/\sigma_1 = 3 \cdot 5$. (ii) The regression line is most markedly skew, in shape like a Galton ogive, so that there is a maximum of regression, and therefore correlation, at the centre of the surface, while the regression and therefore correlation reduce to zero as we move outwards. No frequency surfaces in actual practice exhibit, as far as we are aware, these features demanded by constant Yulean association.

MISCELLANEA.

I. The Correction to be made to the Correlation Ratio for Grouping*.

By STUDENT.

Using the ordinary notation viz. n_{x_ρ} = the number in the x array of y's whose mean is at x_ρ, $\bar{y}_{x_\rho}$ = the mean of this array, N the total number in the sample, and $\bar{y}$ the general mean of y, we have η^2 defined by the relation

$$\eta^2 = \frac{S\{n_{x_\rho}(\bar{y}_{x_\rho} - \bar{y})^2\}}{N\sigma_y^2} \quad \text{.......................................(i).}$$

If η^2 is required to fit a regression curve to the actual observations as in Professor Pearson's original memoir "On the General Theory of Skew Correlation and Non-linear Regression," no correction is necessary.

But if we require a ratio which shall remain constant under wide variations of grouping and of number in the sample and which shall consequently be more comparable from one sample to another, there are two corrections to be made.

The first of these has already been given by Professor Pearson (*Biometrika*, Vol. VIII, p. 256), and he has expressed it as follows:—If $\bar{\eta}^2$ be the value of η^2 actually found by the use of (i), and η^2 be the value which would be found from an infinitely large sample, then if κ be the number of x arrays

$$\eta^2 = \frac{\bar{\eta}^2 - (\kappa - 1)/N}{1 - (\kappa - 2)/N} \quad \text{.......................................(ii).}$$

But there is a further effect of grouping which has not hitherto been noted and which can be evaluated as follows:

Suppose the x_ρ array to be divided into elementary x arrays and let y_p be the mean of the x_p elementary array and n_p its frequency.

Then clearly the proper contribution of the x_ρ array to η^2 is

$$\frac{S\{n_p(y_p - \bar{y})^2\}}{N\sigma_y^2}.$$

This is equal to

$$\frac{S\{n_p(\bar{y}_{x_\rho} - \bar{y} + y_p - \bar{y}_{x_\rho})^2\}}{N\sigma_y^2} = \frac{1}{N\sigma_y^2}\left[S\{n_p(\bar{y}_{x_\rho} - \bar{y})^2\} + 2S\{n_p(\bar{y}_{x_\rho} - \bar{y})(y_p - \bar{y}_{x_\rho})\} + S\{n_p(y_p - \bar{y}_{x_\rho})^2\}\right].$$

Now $\bar{y}_{x_\rho} - \bar{y}$ is of course constant for this summation, $S(n_p) = n_{x_\rho}$ and $S\{n_p(y_p - \bar{y}_{x_\rho})\} = 0$, herefore the contribution to η^2

$$= \frac{n_{x_\rho}(\bar{y}_{x_\rho} - \bar{y})^2}{N\sigma_y^2} + \frac{S\{n_p(y_p - \bar{y}_{x_\rho})\}}{N\sigma_y^2} \quad \text{..............................(iii).}$$

* See above p. 118 of this Journal.

The first of these two terms is that which is obtained in the ordinary way, so the contribution of each array should be corrected by the addition of the second term and η^2 itself by the addition of

$$S\left[\frac{S\{n_p(y_p-\bar{y}_{x_\rho})^2\}}{N\sigma_y^2}\right] \qquad \text{(iv)}.$$

Now if Professor Pearson's correction (ii) has been made we may take the point whose coordinates are (x_p, y_p) to lie on the regression line, and if further we assume the regression line to be linear throughout the x_ρ group and to be inclined at an angle of $\tan^{-1} r_\rho \dfrac{\sigma_y}{\sigma_x}$ to the horizontal we have

$$y_p = x_p \, . \, r_\rho \frac{\sigma_y}{\sigma_x} \text{ and } \bar{y}_{x_\rho} = x_\rho \, . \, r_\rho \frac{\sigma_y}{\sigma_x}.$$

Hence (iv) becomes

$$S\left[\frac{r_\rho^2 S\{n_p(x_p-\bar{x}_\rho)^2\}}{N\sigma_x^2}\right] \qquad \text{(v)}.$$

Now $S\{n_p(x_p-\bar{x}_\rho)^2\}$ is the second moment of the x_ρ group about its own mean and when the distribution is known can often be approximately evaluated. Similarly when the distribution is known r_ρ can be estimated and the correction to η^2 calculated group by group.

But by making certain assumptions we can very much simplify the work, and a practical test, in which the assumptions are not justified, will show the sort of errors which are introduced.

The first assumptions are that the regression is linear and the arrays homoscedastic. In this case of course r_ρ is constant and equal to η; we are practically determining a value of r by the η method.

The correction then becomes

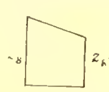
$$\frac{\eta^2}{N\sigma_x^2} S[S\{m_p(x_p-\bar{x}_\rho)^2\}],$$

or writing $\lambda^2 = S[S\{m_p(x_p-\bar{x}_\rho)^2\}]$ and H^2 for the raw value of η^2 after using Pearson's correction, we get from (iii) $\eta^2 = H^2 + \eta^2\lambda^2$ or

$$\eta^2 = \frac{H^2}{(1-\lambda^2)} \qquad \text{(vi)}.$$

To obtain a value for λ^2 we still require to postulate something of the nature of the distribution and I propose to treat (i) of the case where the unit of grouping is constant and small enough for the frequency in each group to be considered to be distributed as a trapezium, and (ii) of the case where the frequency distribution is normal.

(i) First to find the second moment of a trapezium about its mean.



Let z_s and $z_{s'}$ be the ordinates forming the 'walls' of the trapezium and let the group unit be h.

Then $y = z_s + \left(\dfrac{z_{s'}-z_s}{h}\right)x$ is the equation to the 'roof' referred to the 'floor' and left hand 'wall' as axes. The area is clearly $\dfrac{(z_s+z_{s'})\,h}{2}$.

The mean is at

$$\frac{2}{h\,(z_s+z_{s'})}\int_0^h yx\,dx = \frac{2}{h\,(z_s+z_{s'})}\left\{\frac{(z_{s'}-z_s)\,h^3}{3h} + \frac{z_s h^2}{2}\right\} = \frac{h}{3} \cdot \frac{2z_{s'}+z_s}{z_s+z_{s'}}.$$

The second moment coefficient about the axis of y is

$$\frac{2}{h\,(z_s+z_{s'})}\int_0^h yx^2\,dx=\frac{2}{h\,(z_s+z_{s'})}\left\{\frac{(z_{s'}-z_s)\,h^4}{4h}+\frac{z_s h^3}{3}\right\}=\frac{h^2}{6}\cdot\frac{3z_{s'}+z_s}{z_s+z_{s'}}.$$

The second moment coefficient about the mean is

$$\frac{h^2}{6}\cdot\frac{3z_{s'}+z_s}{z_s+z_{s'}}-\frac{h^2}{9}\cdot\frac{(2z_{s'}+z_s)^2}{(z_s+z_{s'})^2}=\frac{h^2}{18}\left\{\frac{z_{s'}{}^2+4z_{s'}z_s+z_s{}^2}{(z_s+z_{s'})^2}\right\}=\frac{h^2}{12}\left\{1-\tfrac{1}{3}\left(\frac{z_s-z_{s'}}{z_s+z_{s'}}\right)^2\right\}.$$

Clearly when h is reasonably small $\left(\frac{z_s-z_{s'}}{z_s+z_{s'}}\right)^2$ is a quantity of the second order and in this case

$$\lambda^2=\frac{h^2}{12\sigma_x{}^2}\quad\text{...(vii)},$$

so that

$$\eta^2=\frac{1}{\left(1-\frac{h^2}{12\sigma_x{}^2}\right)}\left\{\frac{\bar{\eta}^2-\frac{\kappa-1}{N}}{1-\frac{\kappa-2}{N}}\right\}\quad\text{..............................(viii)},$$

when the unit of grouping is uniform and small.

(ii) When the unit of grouping is neither uniform nor small and there is no special knowledge of the nature of the distribution, we must needs fall back on the Gaussian curve to give us a first approximation to z_s and $z_{s'}$ for each group.

In this case

$$1-\lambda^2=NS\left\{\frac{(z_s-z_{s'})^2}{n_{x_p}}\right\}\quad\text{..................................(ix)*},$$

and it is necessary to determine it, after fitting the frequency by means of Sheppard's tables.

Finally, what correction, if any, is to be made for the grouping of y?

This will become more apparent from the alternative formula for η^2, namely

$$\eta^2=1-\frac{S\,(y-\bar{y}_s)^2}{N\sigma_y{}^2}.$$

For the second moment of each array should be corrected by the subtraction of $\frac{n_s k^2}{12}$ where k is the unit of grouping of y so that

$$\eta^2=1-\frac{S\,(y-\bar{y}_s)^2-\frac{Nk^2}{12}}{S\,(y-\bar{y})^2-\frac{Nk^2}{12}}$$

$$=\frac{S\,(y-\bar{y})^2-S\,(y-\bar{y}_s)^2}{S\,(y-\bar{y})^2-\frac{Nk^2}{12}}$$

$$=\frac{S\,(y-\bar{y}_s+\bar{y}_s-\bar{y})^2-S\,(y-\bar{y}_s)^2}{N\sigma_y{}^2}$$

$$=\frac{S(y-\bar{y}_s)^2+2S(y-\bar{y}_s)\,(\bar{y}_s-\bar{y})+S\,(\bar{y}_s-\bar{y})^2-S\,(y-\bar{y}_s)^2}{N\sigma_y{}^2}$$

$$=\frac{S\,\{n_s\,(y_s-\bar{y})^2\}}{N\sigma_y{}^2},$$

since $S\,(\bar{y}_s-\bar{y})^2$ when summed for each individual becomes $S\,\{n_s\,(\bar{y}_s-\bar{y})^2\}$ when summed for each array, and $S\,(y-\bar{y}_s)\,(y_s-\bar{y})$ vanishes for each array.

Hence there is no correction to be made for the y grouping except Sheppard's correction for the Standard Deviation of y.

* The suggestion of this formula I owe to Professor Pearson.

I have tested the results on an instance given in Professor Pearson's original memoir, namely the age and auricular height in Girls, correlation table pp. 34 and 54.

The means of the arrays in the full table are as follows :

Even Grouping Number of Grouping III	II	I	Age	Mean Auricular Height	Number of Cases	Uneven Grouping Number IV	V	VI	VII
		—	*3— 4*	115·25	1				
		—	*4— 5*	116·9643	7				
		—	*5— 6*	117·4722	18				
		—	*6— 7*	119·1000	40				
		—	*7— 8*	120·3026	76				
		—	*8— 9*	121·6340	125				
		—	*9—10*	121·7246	177				
		—	*10—11*	122·8160	235	—			
		—	*11—12*	123·1427	261	—			
		—	*12—13*	123·8908	309	—	—		
		—	*13—14*	124·8622	263	—	—		
		—	*14—15*	125·7146	198	—			
		—	*15—16*	126·1565	214	—			
		—	*16—17*	126·5340	162				
		—	*17—18*	126·9132	95				
		—	*18—19*	127·0205	61				
		—	*19—20*	129·5577	13				
		—	*20—21*	123·8214	7				
		—	*21—22*	126·5000	8				
		—	*22—23*	125·25	2				

These were grouped in seven ways in three of which the groups were of equal width, and the other four give an attempt at equal frequency: the method of grouping is set out by means of columns headed in Roman numerals. The age distribution differs significantly from the normal, the constants being $\beta_1 = \cdot 0013$, $\beta_2 = 2\cdot 7101$, but it would perhaps have been better to have selected a less normal distribution : still it represents the ordinary 'cocked hat' statistics that tend to occur.

The regression is certainly not very linear, the growth apparently ceasing at about 18–19.

The values of $\bar{\eta}^2$ (the raw value), H^2 (the value after using Professor Pearson's correction) and η^2 (the value after attempting to use the λ^2 correction) are given in the following table.

Number of Grouping	Number of Groups	$\bar{\eta}^2$	$\bar{\eta}$	H^2	$\left(\lambda^2 = \frac{h^2}{12\sigma_x^2}\right)$ η^2	η	$1-\lambda^2$ from Normal Curve η^2	η
I	20	·09183	·303	·08414	·08489	·291	·08494	·291
II	10	·08657	·294	·08290	·08595	·293	·08510	·292
III	5	·07701	·278	·07535	·08786	·296	·08635	·294
IV	9	·08836	·297	·08510	—	—	·08953	·299
V	6	·08342	·289	·08136	—	—	·08913	·299
VI	5	·08218	·287	·08053	—	—	·08885	·298
VII	2	·06203	·249	·06159	—	—	·09739	·312

It will be seen that the first three, with even grouping, are very close together though the number of groups has been reduced from 20 to 5. Similarly the next three are close together, and the last is again by itself.

An examination of the way in which the groups are taken shows that the more the tail is bunched together the higher is the value found, and this is what would be expected in this particular case, since there is practically no increase of head height with age at the 'old' end of the scale, whereas for purposes of calculation we have assumed a constant angle for the regression line. But it may be pointed out that η varies (to the 2nd place of decimals) only from ·29 to ·31 even if we reduce the twenty groups to two, an extreme proceeding which is never done in practice.

At the same time the ordinary six or eight groups may be expected to give results a little too high when, as is usual, the regression line is curved.

II. On the Hereditary Character of General Health.

By KARL PEARSON, F.R.S. and ETHEL M. ELDERTON, Galton Laboratory.

(1) In dealing with the heredity of general health we have to meet at once certain fundamental difficulties. We have first the question of environment and secondly the question of variety in health caused by what we may term accident. If we deal with families living in widely differentiated environments we shall have, or certainly may have, a spurious correlation of health in parents and offspring; the resemblance in health will be emphasised. On the other hand, when a single member of a family is exposed to a specially differentiated environment, i.e. goes to the West Coast of Africa, or spends his life in India, or catches enteric at a particularly unfavourable moment, the correlation of general health may be decidedly weakened in the case of parent and offspring. These difficulties of differentiated environment and what we may, perhaps, term accident cannot be wholly overcome, but we may endeavour to meet or measure them. In the first place we can confine our observations to one social class and thus go a long way to get differentiated environment removed. If, as in the present paper, we deal essentially with the professional classes, there is great uniformity of general environment. The food supply is sufficient, the doctor is always at command, physical exercise is fairly general and markedly insanitary houses or occupations are practically avoided. We do not think therefore that, for the data of the present paper, differential environment is a marked factor in producing correlation. On the other hand we do consider it possible that "accident" will weaken the relationships sought. The reduction in health-correlations below the values for other physical characters, might indeed be taken as a measure of random action on health, comparable with the random action of death itself in reducing the correlation of duration of life, which has already been discussed by one of us*. Indeed heredity of general health is almost as significant for the problem of natural selection, as heredity of duration of life.

A more serious difficulty in the health-inheritance problem is this very question of death. If parents are delicate and health or delicacy is hereditary, they will have delicate children, and we may anticipate that more of these children will die than in the case of the children of robust or normally healthy parents. Thus only the healthier children of delicate parents will survive for us to record their state of health, and accordingly the offspring of delicate parents will appear healthier than they really should be owing to the selective death-rate. In our investigations we have dealt only with offspring who lived to be adult, i.e. to at least 21 years of age, so that an appreciation of general health could be formed. There is also a further difficulty that very delicate parents are themselves likely to die, and we have again

* *R. S. Proc.* Vol. 65, p. 290, and *Biometrika*, Vol. I. p. 50.

less chance of getting their record. Hence the problem is by no means without difficulty, but it is of such great interest that we venture to give here our chief results.

Our health classes were: Very Robust, Robust, Normally Healthy, Rather Delicate, Delicate and Very Delicate. These have been verbally defined, and our present definitions are as follows:

SCALE.

V.R. Very robust.—He has never had to see a doctor, nor been off work through illness.

R. Robust.—He has only seen a doctor about minor ailments, and has only been off work for colds, etc.

N.H. Normally healthy.—He has not had more than one serious illness, involving, say, a fortnight's absence from work during the last ten years.

R.D. Rather delicate.—He has had more than one serious illness, but not more than one involving more than four weeks' absence from work during the last ten years.

D. Delicate.—He is off work through illness at least four weeks in all every year.

V.D. Very delicate.—He is in a chronic state of ill-health.

But these definitions were not used throughout the whole of the records included in the present investigation, and we found considerable reluctance to the use of the "Very Delicate" category. Accordingly in the present reduction all the delicate categories have been clubbed together. As in previous investigations a Gaussian frequency scale was used for Health*, the interval on the scale covered by "Normally Healthy" being taken to represent 100 units of health, which may be called *sanitaces*. These hundred sanitaces were supposed to represent the range of normal health of each type of individuals—fathers, mothers, sons and daughters, and the rest of the distribution calculated in terms of them. At the mean of the parental category was then plotted up the mean number of sanitaces of all the offspring of either sex of parents of the given category. Thus the round black dots of our four first diagrams were obtained. It was found that the three points marking the mean health of offspring of "Very Robust," "Robust" and "Normally Healthy" parents were in all four cases closely on a line which passed through the mean of parental and filial health, but that the mean health of children of delicate parents in all four cases gave a point lying markedly above this line. This line in all four cases shows a marked slope indicating that as the health of the parents is worse so the health of their offspring is worse.

To illustrate the probable source, or at least part source, of the anomaly in the case of the delicate parents, the percentage of sons dying before 21 years of age and so escaping record of health was obtained for each health group of the fathers and this percentage was plotted in Diagram V. to a Gaussian scale of Father's health. Again we note a fairly uniform increase of the percentage of deaths with decreasing goodness of health of the father, for the three better categories, but a very marked excess of deaths—quite off the line—in the category of delicate fathers, where some 13 °/₀ of sons die as against the 7 °/₀ in the case of "Very Robust" fathers.

(2) Still preserving our Gaussian scales, we may consider that it is the sons of delicate fathers who have been especially selected. In this case the health of the sons of the three left-hand categories of fathers will give very nearly the true regression line for unselected material. We have made no attempt to algebraically fit the best lines to these three points, but placed them graphically on the points and through the general means†.

* Cf. Huxley Lecture, *Biometrika*, Vol. III. p. 146.

† This was done because the three left-hand points always fall nearly on the same line with the general means, but of course the means of both parents and offspring would be somewhat lowered had we the full complement of delicate sons of delicate parents:—to increase the number of such sons would *ipso facto* increase the number of such parents and thus both means of health in offspring and parents would be lowered in the same direction probably approximately that of the true regression line.

But there is another disturbing factor, which the diagrams make at once obvious, the mean health of parents is in very considerable degree better than that of the offspring. Measuring our mean health of any population in sanitaces from the division between "Robust" and "Normally Healthy," *negative* towards normally healthy, we have the following results :

TABLE I.

Mean Healths of Various Populations.

		Δ
Fathers of Sons :	+ 2·72 sa.	+4·81.
Fathers of Daughters :	− 2·09 sa.	
Mothers of Sons :	−20·22 sa.	−2·49.
Mothers of Daughters :	−17·73 sa.	
Sons of Fathers :	−17·11 sa.	−0·29.
Sons of Mothers :	−16·82 sa.	
Daughters of Fathers :	−35·53 sa.	−0·89.
Daughters of Mothers :	−34·64 sa.	

These results show at once that the health of the parents is far better than that of their offspring. This does not imply that the younger generation has degenerated but only that there is a selection of the more healthy for parentage; the more robust men and women are, the more likely they are to be parents and repeated parents. The differences between sons of fathers and sons of mothers is due to a difference of material and the same applies to daughters of fathers and daughters of mothers; these differences are probably those of random sampling. The differences in the cases of parents are more marked and possibly significant. If they be, then we have some slight suggestion that a healthy father and a delicate mother would be more likely to have sons and a delicate father and a healthy mother daughters. Even if there be anything in the suggestion, it would only be shown in large numbers, and is not a universal rule for individual pairs. All we can say is that the numbers do not flatly contradict a popular impression of the kind.

(3) Another result brought out by our numbers is that the health of the male is markedly better than that of the female in both generations; this is possibly the effect of a more stringent selection of the male. Such a selection may be of two kinds, first the known heavier death-rate of the male, and secondly a greater objection to admission of delicacy or even to giving a record at all on the part of the delicate male. The general difference can be seen in the following percentages of delicate individuals :

TABLE II.

Percentages of Delicate Individuals in Various Populations.

Fathers of Sons :	7·76 °/₀.
Fathers of Daughters :	9·68 °/₀.
Sons of Fathers :	13·69 °/₀.
Sons of Mothers :	13·57 °/₀.
Mothers of Sons :	16·65 °/₀.
Mothers of Daughters :	16·56 °/₀.
Daughters of Fathers :	20·44 °/₀.
Daughters of Mothers :	20·82 °/₀.

There is an increase in the delicacy rate as we pass from Fathers of Sons to Fathers of Daughters; a slight fall only in the delicacy rate as we pass from Mothers of Sons to Mothers of Daughters. But both are supported by the rates of "Very Robust" where we find the following percentages: Fathers of Sons 18·76; Fathers of Daughters 17·91; Mothers of Sons 11·28; Mothers of Daughters 13·02.

DIAGRAM I.

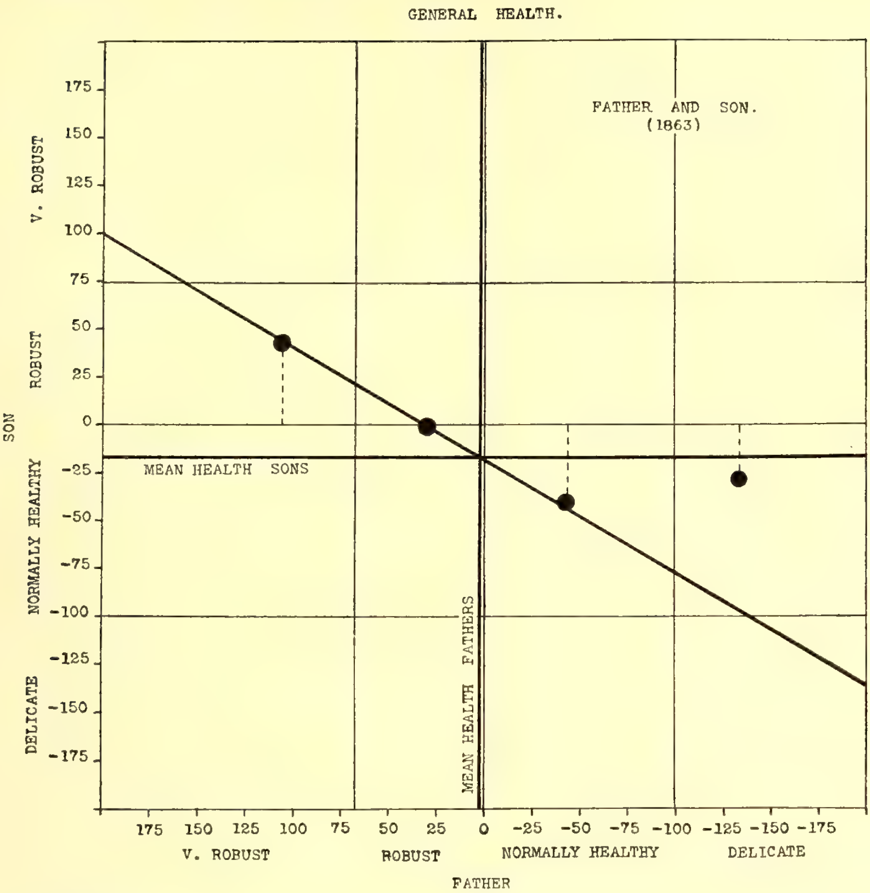


DIAGRAM II.

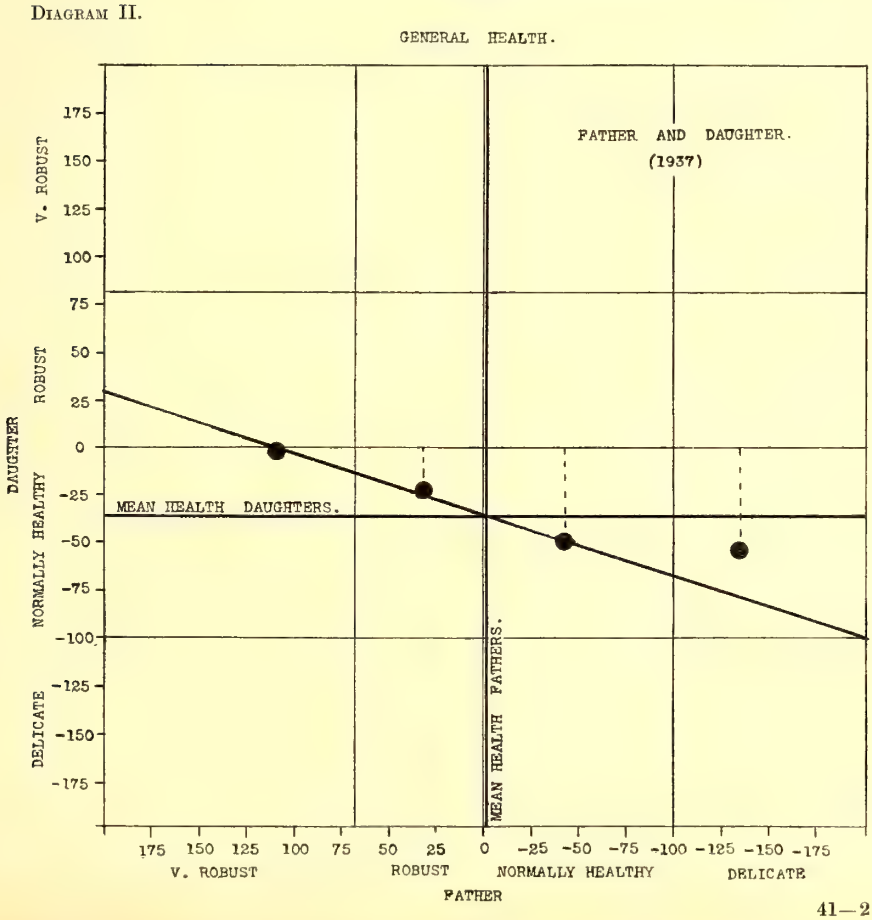


Diagram III.

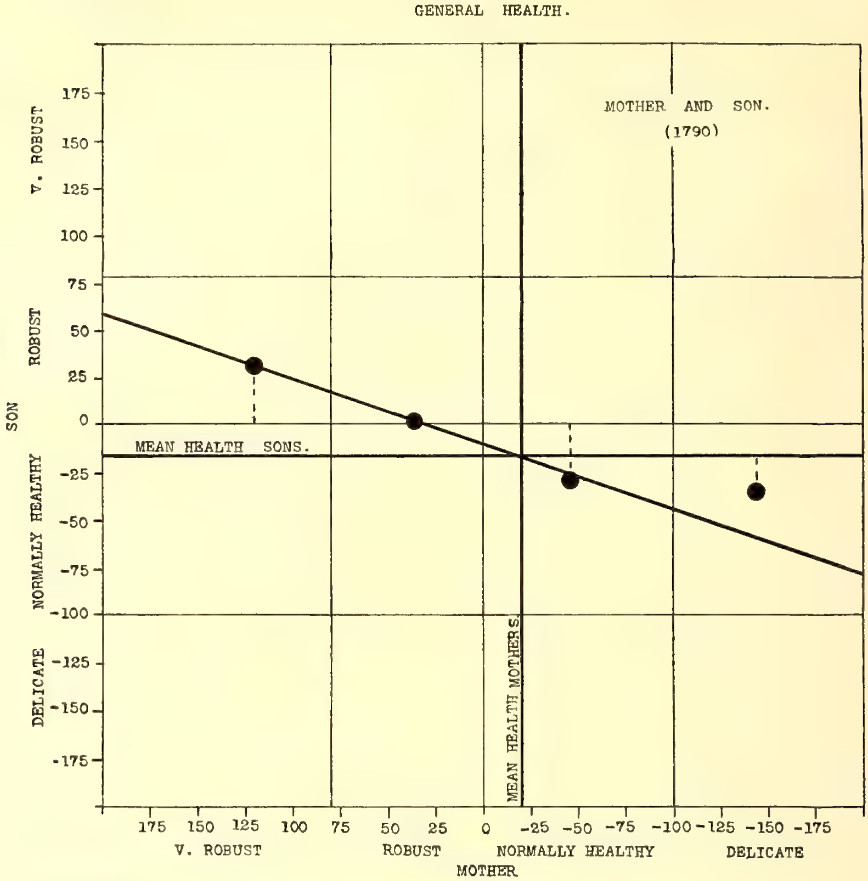


Diagram IV.

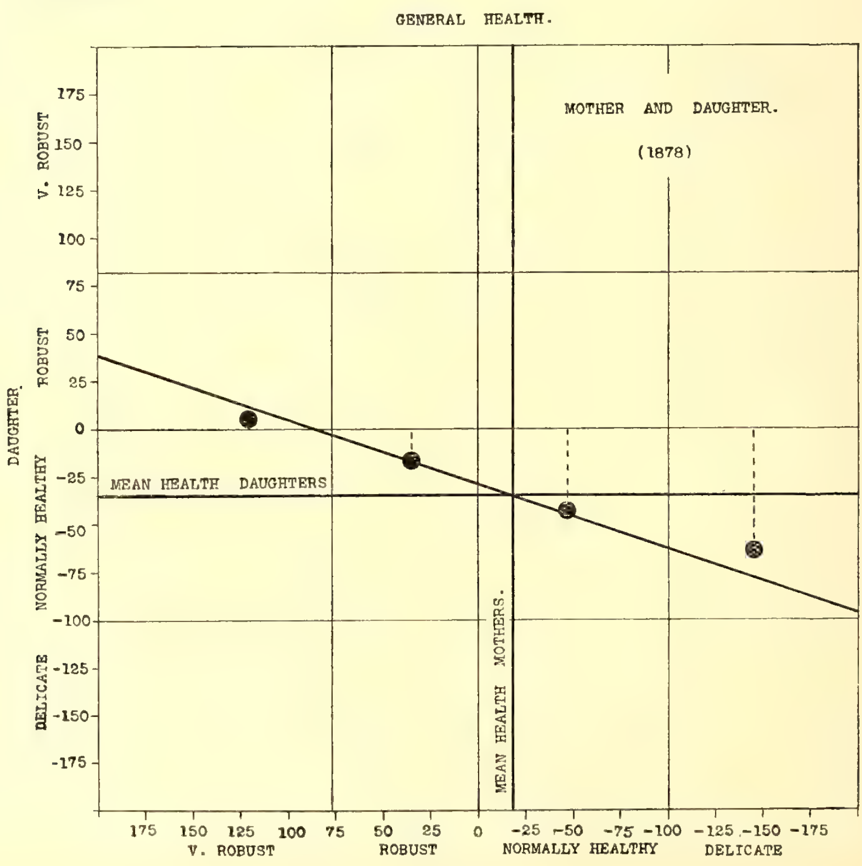
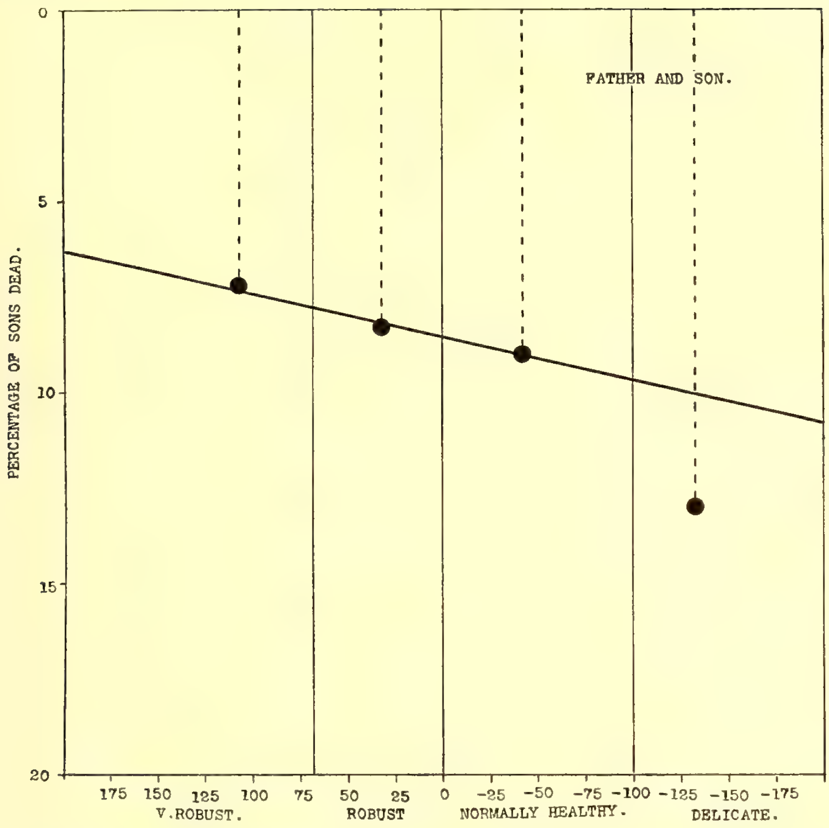


Diagram V.



(4) In the following table the Standard Deviations of the various populations are given also in terms of sanitaces.

TABLE III.

Standard Deviations of Health in Terms of Sanitaces of Various Populations.

Fathers of Sons :	72·25 sa.
Fathers of Daughters :	75·31 sa.
Mothers of Sons :	82·40 sa.
Mothers of Daughters :	84·65 sa.
Sons of Fathers :	75·73 sa.
Sons of Mothers :	75·64 sa.
Daughters of Fathers :	78·06 sa.
Daughters of Mothers :	80·42 sa.

The definite conclusions we can draw from this table are : first, that women are definitely more variable in health than men, and secondly that daughters seem less variable in health than their mothers, while sons are possibly but only slightly more variable than their fathers. On the whole the nature of the health selection between the older and younger generations seems to be of a character which leaves the variability only slightly modified, but shifts very considerably the mean health as a whole from parents to children.

(5) The results so far reached enable us to test the degree of stability of our scale. If that scale is a reasonable one, the range of the "Robust" category ought to come out with reasonable sameness when measured in sanitaces for the different populations. We find:

TABLE IV.

Range of Robust for Various Populations.

Fathers of Sons:	66·80 sa.
Fathers of Daughters:	67·02 sa.
Sons of Fathers:	74·80 sa.
Sons of Mothers:	76·95 sa.
Mothers of Sons:	79·61 sa.
Mothers of Daughters:	77·52 sa.
Daughters of Fathers:	81·54 sa.
Daughters of Mothers:	82·47 sa.

It will, we think, be clear from these results that the terms employed in our categories have been used in rather different senses when applied to males and females, and when applied to the younger and older generation. If we assume "Normal Health" to be the same for all types, then the category "Robust" has been used in a wider sense for women than for men, and in a wider sense for the younger than the older generation. Nor does this seem unreasonable when we compare ordinary practice, in which undoubtedly a different health scale is applied to men and women, and to old and young. As we have seen there is on an average a lower state of health in women than in men and we are apt to judge by deviation from the average rather than by absolute condition. Again the average health of the older generation is higher than that of the younger, and one is rather apt to compare the health of the offspring with that of the parent instead of applying an absolute standard to both. Anyhow without laying much stress on the reasons for the personal equation, it is probable that judgment does differ in the matter of health according as we are dealing with man or woman, and with the old or young generation.

Speaking in quite round numbers we may say that the range of our "Robust" in women, regardless of whether they are parents or not, is about 80 sanitaces and this is almost equal to the standard deviation of their health (mean of four populations 81·4 sa.). In the case of men of every status, the range (71·39 sa.) of the "Robust" is slightly less than the standard deviation (74·7 sa.), but for many purposes it may be sufficient to consider both as 75 sanitaces. We have no means of ascertaining the absolute health in sanitaces of any individual, but if we were to assume 300 sanitaces as the stock of an individual on the border of "Normal Health" and "Robustness," the average man would have about 292 and the average woman about 273 units of health; one woman in a thousand would have less than 29 units of health, and one man less than 67. The "Very Robust" man would be a man with more than 371 units and the "Very Robust" woman would be a woman with more than 380 units. Finally one man and one woman also in a thousand would have more than 517 units of health. Thus while the most robust men and women in the thousand are of the same calibre, the most delicate woman has less health than the most delicate man,—a result possibly of the more stringent death-rate; a man needs more health to survive at all. Of course these results are purely suggestive, but they flow with some probability from the lower average health of women and their greater variability. We should not desire to place any great weight on them. Other data for different age groups and social classes will be discussed later, and then it will be more possible to propose with greater certitude a definite health scale.

(6) Looked at from the average environment of the professional classes, there can be little doubt that our diagrams indicate that general health is a hereditary character; but we have of

course to regard the difficulty of the record in regard to delicate offspring of delicate parents. We want if possible to get some measure of the correlation in health of parent and offspring from our first three points in each diagram.

Clearly we can find the slopes of the regression line passing through these three points. If σ_O be the standard deviation of offspring and σ_P of parents, the slope

$$s = \frac{\sigma_O}{\sigma_P} \times r_{OP}.$$

This of course is independent of normality. Again the weighted mean square deviation of arrays has the value $\sigma_O^2(1-r_{OP}^2)$. Let the S. D.'s of Very Robust, Robust, and Normally Healthy be σ_{a_1}, σ_{a_2}, and σ_{a_3}, then

$$\sigma_O^2(1-r_{OP}^2) = \frac{n_1\sigma^2_{a_1} + n_2\sigma^2_{a_2} + n_3\sigma^2_{a_3}}{n_1+n_2+n_3},$$

$$\therefore \frac{1-r_{OP}^2}{r_{OP}^2} = \frac{n_1\sigma^2_{a_1} + n_2\sigma^2_{a_2} + n_3\sigma^2_{a_3}}{(n_1+n_2+n_3)\,\sigma_P^2 \times s^2} \,\cdots\cdots\cdots$$

Now σ_{a_1}, σ_{a_2}, σ_{a_3} can for each array be expressed in terms of h the 100 sanitaces of the "Normally Healthy" range for that array and s can be carefully measured on the diagrams. It only remains to consider what value shall be given to σ_P. Undoubtedly some parents are omitted because they have died from delicacy, but on the whole we are convinced that there has been rather less selection of parents than of offspring. Accordingly we have put σ_P its value in terms of the range of "Normally Healthy." Thus r_{OP} can be calculated, without regarding the final anomalous array.

But clearly we have to correct the result for our grouping in arrays of parents, but for parents only, as the S. D.'s of the arrays have been found from total frequencies of the groups, on the assumption that each array is normal. We must therefore divide each correlation by the correlation between class-index and individual character—a point discussed in another paper (see pp. 116 and 134 above).

These corrective factors, r_{xC_x} of our notation, are:

r_{xC_x} for Fathers in	Fathers and Sons:	·9258.
„ „	Fathers and Daughters:	·9333.
„ Mothers in	Mothers and Sons:	·9354.
„ „	Mothers and Daughters:	·9384.

Thus we have

Correlations of Health, Parent and Offspring.

	Raw*	Corrected	
Father and Son:	·4456	·4813	·39.
Father and Daughter:	·2852	·3056	
Mother and Son:	·3551	·3796	·37.
Mother and Daughter:	·3407	·3631	

Mean = ·3824.

We have also considered the correlations from another standpoint. We have for the slope s of the regression line for the three first points

$$s = \frac{\sigma_O}{\sigma_P} r_{OP}.$$

* Corrected of course for defect of delicate offspring of delicate parents, i.e. found from formula for r_{OP} above.

Therefore :

$$r_{OP} = \frac{\sigma_P}{\sigma_O} \times s.$$

Now σ_P may be slightly too small because some delicate parents will be omitted and σ_O will probably be too small because, as we have seen, many delicate offspring escape record. We shall thus get rather too large values of r_{OP}. There will be no correction to be made this time as the values of σ_P and σ_O are based on frequencies occurring between certain limits. The following values were obtained :

	Slope s	Correlation	
Fathers and Sons :	·55	·5247	·42.
Fathers and Daughters :	·33	·3184	
Mothers and Sons :	·345	·3758	·37.
Mothers and Daughters :	·34	·3578	

Mean = ·3942.

This confirms the values previously obtained and we think we may adopt them and especially the mean* value ·3824 first reached as at least a fair approximation. Now suppose the chance of an individual's health being due to some other cause than heredity to be p, then in a population of N pairs the chance of both pairs having their health as a natural inheritance will be $(1-p)^2$ and the ratio of correlated to the total material will be $(1-p)^2 N/N$, this will measure the reduction in the correlation of health between parent and offspring due to accidental and extraneous causes.

Now the full strength of heredity for physical characters has been shown in the professional classes to be on an average about ·46 : see *Biometrika*, Vol. II. p. 357. Hence we have :

$$(1-p)^2 = \cdot 3824/\cdot 46,$$

or :

$$p = \cdot 0743.$$

Thus the health of only 7·4 °/。 of the population is determined by accidental causes. In 92·6 °/。, i.e. in the great bulk of cases, heredity is the chief factor in the determination of general health. Without laying too much stress on the actual figures, we think it may safely be concluded that heredity in the professional classes is the chief source of the good or bad health of individuals.

Tables. General Health.

TABLE I.

Father's Health.

Son's Health.	Very Robust	Robust	Normally Healthy	Delicate	Totals
Very Robust	111·25	63·75	19	15·5	209·5
Robust	117·25	237·5	161·5	39·25	555·5
Normally Healthy ...	88	236	461·5	57·5	843
Delicate	33	72·75	117	32·25	255
Totals	349·5	610	759	144·5	1863

* There is something anomalous in the high value of the father and son correlation, but we have not been able to trace it to any definite origin.

TABLE II.

Mother's Health.

Son's Health.	Very Robust	Robust	Normally Healthy	Delicate	Totals
Very Robust	76	47·5	42	27	192·5
Robust	55·5	215·75	201·5	72·25	545
Normally Healthy ...	61	189·75	424·5	134·25	809·5
Delicate	9·5	66·5	102·5	64·5	243
Totals	202	519·5	770·5	298	1790

TABLE III.

Father's Health.

Daughter's Health.	Very Robust	Robust	Normally Healthy	Delicate	Totals
Very Robust	53·5	33·25	25·5	16·25	128·5
Robust	117·5	207	149	26·5	500
Normally Healthy ...	110	245·25	462·25	95	912·5
Delicate	66·5	114	165·75	49·75	396
Totals	347·5	599·5	802·5	187·5	1937

TABLE IV.

Mother's Health.

Daughter's Health.	Very Robust	Robust	Normally Healthy	Delicate	Totals
Very Robust	52	38·5	30·5	15·5	136·5
Robust	76	191·25	164·25	58	489·5
Normally Healthy ...	85·5	213	431·75	130·75	861
Delicate	31	95·75	157·5	106·75	391
Totals	244·5	538·5	784	311	1878

III. Note on the Honduras Piebald.

By KARL PEARSON, F.R.S.

The first report of this piebald with photographs was, I believe, brought back to Europe by M. le Comte Maurice de Périgny. Professeur R. Blanchard published an account based on the Comte de Périgny's data in the *Bulletin de la Société française d'Histoire de la Médecine*, t. IX. p. 213, Paris, 1910. Through the courtesy of the Comte de Périgny I have been provided with copies of his photographs, which will appear in the Second Part of the *Monograph on Albinism* soon to be issued by E. Nettleship, C. H. Usher and myself, and are reproduced as Plates X. and XI. here. This piebald boy is of much interest because he belongs essentially to the "classical type," illustrated in paintings of the 18th century, and no living piebald of this type had so far come to our knowledge, although we had in the First Part of our monograph given many illustrations of early cases.

But there is further scientific interest in this piebald because he supports the point of view emphasised by me in the monograph referred to, that when a pure race is crossed with a mixed race—a hybrid between races with markedly different degrees of pigmentation—then piebalds are likely to appear *de novo*. On this ground it seems to me idle to speak of piebaldism as a Mendelian unit character, and it is idle equally to talk of it as a latent unit character, for there is no evidence at all that it ever occurred before in the pure races whose crossing leads to these piebalds. In the present case the mother has Mexican and negro blood, the Mexican being already a mixture of Spanish and American Indian. She thus combines black, red and white races. The father is a coal-black pure African. But several of our piebald pedigrees show that it is sufficient for piebaldism when a pure red or white is crossed with a pure dark race and then the hybrid be mated again with either pure race, the produce of this second cross may be a piebald. The light race may be merely an albino variety of the dark race, and in several of our dark race pedigrees we find such piebalds occurring in stocks wherein albinism has occurred also. Indeed given an albino occurring in a dark race it seems possible from the hybrid between normal and albino by crossing again with normal or albino to produce almost every shade of colour as well as every variety of piebaldism. If such cases occur in man and axolotl, it seems unnecessary to seek in the ancestry of the albino for possibilities of either piebaldism or colour, as for example has been done in the case of mice. There is no reason to suppose brilliant colours latent in the normal axolotl, nor piebaldism latent in the normal man. At any rate where there is no evidence of it before the crosses take place, it is more reasonable to suppose it a product of the crosses themselves.

Without entering fully into a matter which will shortly be discussed at length elsewhere, I would point out that when a hybrid is formed between a black dog which has bred true for many generations and an albino of another race, this hybrid is either black or black with white markings on chest, never in our experience so far a true piebald, but when this hybrid is crossed again with the albino, we obtain at once not only black dogs, or black dogs with white markings, but black and white piebalds, lilac and white piebalds, rusty black dogs, red dogs, and albinos; possibly as the work goes on other types will appear also. Much the same changes seem to arise from like crosses not only in mice and axolotl but also in man.

In the case of our present piebald we have the father a pure coal-black negro and his brothers and sisters are like him, the mother is a mixture of negro and Spanish-Indian blood; she has a fair skin, black eyes and long black hair, and brothers and sisters are fair skinned like her. There are six children; the three eldest are boys aged 11 years, 9 years, and 7½ years

Lisbey, the Piebald Boy of El Cayo, British Honduras. From a photograph taken for K. Pearson, through the courtesy of Robert H. Franklin, Esq., July, 1912. Lisbey's Father, Mother, and three younger sisters are shown.

Lisbey, the Piebald Boy of El Cayo, British Honduras. From a photograph taken for K. Pearson, through the courtesy of Robert H. Franklin, Esq., July, 1912. Plates VIII and IX compared with X and XI indicate that the frequency and relative size of the colour patches appear to have altered.

Lisbey, the Piebald Boy of El Cayo, British Honduras. From a photograph most kindly provided by M. le Comte Maurice de Périgny. Lisbey is shown with his Father. Front view, 1908. Reproduced from Pearson, Nettleship and Usher's *Monograph on Albinism*, Part II.

Lisbey, the Piebald Boy of El Cayo, British Honduras. From a photograph most kindly provided by M. le Comte Maurice de Périgny. Lisbey is shown with his Father. Back view, 1908. Reproduced from Pearson, Nettleship and Usher's *Monograph on Albinism*, Part II.

respectively, the last being the piebald; the three youngest are girls aged 4 years, 2 years, and 9 months respectively. These last three children together with the piebald are given in Plate VIII. The grandparents were all "normal," i.e. not piebald. Plate VIII. suffices to show not only the difference in colour between father and mother, but the range of colour in the three daughters.

As soon as I had seen M. Blanchard's account of this piebald boy I wrote to the District Commissioner, at El Cayo, Mr R. H. Franklin, and he most kindly sent me the fuller particulars here given, as well as arranged to have the family photographed for me. He tells me that the youngster is the pet of the place as well as its "curiosity" and that he is quite an intelligent boy. He was born on December 1st, 1904, at Peten in Guatemala and the reason his parents give for his piebaldism is "that owing to an eclipse of the moon on the night of his birth, he caught it in the head and it scattered over his body." There was no eclipse of the moon on that date nor any near it.

One point further may be emphasised with regard to this boy. In our Monograph on Albinism, p. 248, Plate I (23)—(26), we deal with the case of a piebald boy from Papua and we give photographs taken at nine years interval. The dark patches have grown larger, but they have not increased in number nor in relative size. In the present case a careful comparison of the dark patches on the legs and arms of this Honduras piebald boy in the recent photographs (Plates VIII. and IX.) of July 1912 and those of the Comte de Périgny of some four years earlier (Plates X. and XI.), shows with almost absolute certainty that the relative area of dark pigmentation has considerably *increased*. This remarkable fact renders the boy of special interest, and it is to be hoped that in still later years photographs of him may be obtainable for comparison with the present series.

IV. Selection and Intermediates in Bacillus coli.

By LEONARD KEENE HIRSHBERG, M.A., M.D.

In the course of some other work on a strain of Bacillus coli, taken from the rectum of a Scotch collie, and planted first in beef tea on October 5th, 1911, transplantations and agar plates were made of these organisms with the original purpose of studying what has been hitherto called involution forms. The course of this work directed my attention to the possibility of these forms being dependent upon the quantity and quality of the pabulum or nutrient material furnished to the bacteria, and hence these so-called involution forms being actually types of polymorphism.

Incidentally observations were made, in view of the claims made by the biologists represented by Professor Castle of Harvard on the one side, and Professor H. S. Jennings on the other, along the lines of possibly selecting races of long, short, and various intermediate generations of this colon bacillus. If it had proved possible to select a type of long, intermediate, or short bacilli, that had remained within the limits of the select mode without reverting back or possessing the power to generate all the types, it would have strengthened the work of Professor Castle on such higher animals as rats, with their manifold and necessarily complexly interrelated factors. As it is, however, after making two hundred and twenty transplantations of colonies of this strain of organisms, and trying to generate true long, short, thin, narrow, and intermediate types for both length and thickness, I find at the conclusion of that part of the work that there is absolutely no ground in my experiments supporting selection as an element in generating a particular type of these bacteria.

Placed in suitable nutrient media, at room temperature, these bacilli divide very rapidly by simple fission. From twenty minutes to half an hour is the average time of division, yet if we

allow an hour, it is seen that in twenty-four hours—the average time of the transplantations on tubes or plates of various media for selecting—one bacillus gives generations numbering seventeen millions of separate individuals.

When the nutrient material is favourable and incubator temperature is used, the types of the bacilli tend to be shorter and actually thinner—although relatively thicker in appearance—while the rate of fission is much increased. In solid media there is also a greater tendency to the small forms, while in peptone and other liquid and at the same time less favourable material the long tenuous and slowly dividing types are more common.

When bizarre types like those formerly called "degenerated bacilli" such as flask-shaped, drum-stick, dumb-bell, Indian club, and tenpin-like variations made their appearance, it was always possible to transplant these and obtain polymorphous bacilli of all the previously known specimens.

True enough, it was often difficult at first to start these unfavourably shaped types growing. They required close attention, frequent transplantations, and the best media such as milk, serum bouillon, and sugar agar or sugar gelatin, but they inevitably came around, and again generated every type that had been observed during the experiments.

Hanging drop slides as well as basic aniline dyes were used in the work, and although there was no standard speed by which the motility factor could be studied, an incidental attempt was made to select slowly moving from rapidly motile organisms. This too was without success, and also depended evidently on food supply, temperature, and other environmental changes. The succeeding generations always produced both types and none generated true forms of motility.

Although some attempt was made to call all types under two micromillimetres short, and all over six micromillimetres long, the intermediates were allowed a range of three to five micromillimetres. All above five-tenths of a micromillimetre were thick, while thin ones were one or two-tenths.

Summary and conclusion: From these experiments, which it must be emphasized are incidental to some other bacteriological studies, it seems that in the case at least of Bacillus coli, a condition of polymorphism exists.

Efforts at selection in two hundred and twenty-five transplantations of thousands of generations each, resulted in absolute failure to obtain any true strain of form or motility. The organisms, while subject to great variations about the given mode of the variety according to its food and environment, always reverted to the previous mean in subsequent generations.

All types were under the proper conditions possessed of the power of generating all other types, hence selection as a method of generating any of these or any new type brought no result.

CORRIGENDA.

In Vol. VIII. pp. 262—6 in the "Study of Pygmy Crania" by Miss H. Dorothy Smith, a slip occurs which is several times repeated. The crania dealt with are described as of the Third Dynasty. They belong to the XXVI—XXX Dynasties. As to the cemetery from which they were taken: see Flinders Petrie, *Gizeh and Rifeh*, p. 29.

Mr J. I. Craig wishes to state that he has discovered the difference between Professor Myers' measurements of head-length and those used in his own paper on the "Anthropometry of Modern Egyptians" (*Biometrika*, Vol. VIII. p. 78). Professor Myers measured as usual from the glabella; Mr Craig was told that the prisoners' heads were measured in the usual way (see *loc. cit.* p. 67, § 4), and supposed this to be also from the glabella. But he now finds that the Egyptian criminals are measured from the *nasion*. This explains the large differences between the head-lengths of the criminals and soldiers commented on in the Editorial footnote on p. 78.

CONTENTS

The publication of a paper in *Biometrika* marks that in the Editor's opinion it contains either in method or material something of interest to biometricians. But the Editor desires it to be distinctly understood that such publication does not mark assent to the arguments used or to the conclusions drawn in the paper.

Biometrika appears about four times a year. A volume containing about 500 pages, with plates and tables, is issued annually.

Papers for publication and books and offprints for notice should be sent to Professor KARL PEARSON, University College, London. It is very desirable that a copy of all measurements made, not necessarily for publication, should accompany each manuscript. In all cases the papers themselves should contain not only the calculated constants, but the distributions from which they have been deduced. Diagrams and drawings should be sent in a state suitable for direct photographic reproduction, and if on decimal paper it should be *blue* ruled, and the lettering only pencilled.

Papers will be accepted in German, French or Italian. In the first case the manuscript should be in Roman not German characters.

Contributors receive 25 copies of their papers free. Twenty-five or fifty additional copies may be had on payment of 3/6 or 6/- respectively per sheet of eight pages; these should be ordered when the final proof is returned.

The subscription price, payable in advance, is 30*s*. *net* per volume (post free); single numbers 10*s*. *net*. Volumes I, II, III, IV, V, VI, VII and VIII (1902—12) complete, 30*s*. *net* per volume. Bound in Buckram 34/6 *net* per volume. Index to Volumes I to V, 2*s*. *net*. Subscriptions may be sent to C. F. Clay, Cambridge University Press, Fetter Lane, London, E.C., either direct or through any bookseller, and communications respecting advertisements should also be addressed to C. F. Clay.

Till further notice, new subscribers to *Biometrika* may obtain Vols. I—VIII together for £9 *net*—or bound in Buckram for £11 *net*.

CAMBRIDGE: PRINTED BY JOHN CLAY, M.A. AT THE UNIVERSITY PRESS.

Vol. IX. Parts III and IV. October 1913

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

CAMBRIDGE UNIVERSITY PRESS
C. F. CLAY, MANAGER
LONDON: FETTER LANE, E.C.
EDINBURGH: 100, PRINCES STREET
also
H. K. LEWIS, 136, GOWER STREET, LONDON, W.C.
WILLIAM WESLEY AND SON, 28, ESSEX STREET, LONDON, W.C.
CHICAGO: UNIVERSITY OF CHICAGO PRESS
BERLIN: A. ASHER AND CO.
LEIPSIC: BROCKHAUS
BOMBAY AND CALCUTTA: MACMILLAN AND CO., LIMITED
TORONTO: J. M. DENT AND SONS, LIMITED
TOKYO: THE MARUZEN-KABUSHIKI-KAISHA
Entered at the New York Post Office as Second Class Matter

Price Twenty Shillings net.

[*Issued November* 8, 1913]

BIOMETRIC LABORATORY PUBLICATIONS

Drapers' Company Research Memoirs.

Biometric Series.

I. **Mathematical Contributions to the** Theory of Evolution.—XIII. On the Theory of Contingency and its Relation to Association and Normal Correlation. By KARL PEARSON, F.R.S. *Issued.* Price 4*s. net.*

II. **Mathematical Contributions to the** Theory of Evolution.—XIV. On the Theory of Skew Correlation and Non-linear Regression. By KARL PEARSON, F.R.S. *Issued.* Price 5*s. net.*

III. **Mathematical Contributions to the** Theory of Evolution.—XV. On the Mathematical Theory of Random Migration. By KARL PEARSON, F.R.S., with the assistance of JOHN BLAKEMAN, M.Sc. *Issued.* Price 5*s. net.*

IV. **Mathematical Contributions to the** Theory of Evolution.—XVI. On Further Methods of Measuring Correlation. By KARL PEARSON, F.R.S. *Issued.* Price 4*s. net.*

V. **Mathematical Contributions to the** Theory of Evolution.—XVII. On Homotyposis in the Animal Kingdom. By ERNEST WARREN, D.Sc., ALICE LEE, D.Sc., EDNA LEA-SMITH, MARION RADFORD, and KARL PEARSON, F.R.S. *Shortly.*

VI. **Albinism in Man.** By KARL PEARSON, E. NETTLESHIP, and C. H. USHER. Text, Part I, and Atlas, Part I. *Issued.* Price 35*s. net.*

VII. **Mathematical Contributions to the** Theory of Evolution.—XVIII. On a Novel Method of Regarding the Association of two Variates classed solely in Alternative Categories. By KARL PEARSON, F.R.S. *Issued.* Price 4*s. net.*

VIII. **Albinism in Man.** By KARL PEARSON, E. NETTLESHIP, and C. H. USHER. Text, Part II, and Atlas, Part II. *Issued.* Price 30*s. net.*

IX. **Albinism in Man.** By KARL PEARSON, E. NETTLESHIP, and C. H. USHER. Text, Part IV, and Atlas, Part IV. *Issued.* Price 21*s. net.*

Studies in National Deterioration. Price 3*s. net* each.

I. **On the Relation of Fertility in Man** to Social Status, and on the changes in this Relation that have taken place in the last 50 years. By DAVID HERON, M.A., D.Sc. *Issued.* Sold only with complete sets.

II. **A First Study of the Statistics of** Pulmonary Tuberculosis (Inheritance). By KARL PEARSON, F.R.S. *Issued.*

III. **A Second Study of the Statistics of** Pulmonary Tuberculosis. Marital Infection. By ERNEST G. POPE, revised by KARL PEARSON, F.R.S. With an Appendix on Assortative Mating by ETHEL M. ELDERTON. *Issued.*

IV. **The Health of the School-Child in** relation to its Mental Characters. By KARL PEARSON, F.R.S. *Shortly.*

V. **On the Inheritance of the Diathesis** of Phthisis and Insanity. A Statistical Study based upon the Family History of 1,500 Criminals. By CHARLES GORING, M.D., B.Sc. *Issued.*

VI. **A Third Study of the Statistics of** Pulmonary Tuberculosis. The Mortality of the Tuberculous and Sanatorium Treatment. By W. P. ELDERTON, F.I.A. and S. J. PERRY, A.I.A. *Issued.*

VII. **On the Intensity of Natural Selection** in Man. By E. C. SNOW, D.Sc. *Issued.*

VIII. **A Fourth Study of the Statistics of** Pulmonary Tuberculosis: the Mortality of the Tuberculous: Sanatorium and Tuberculin Treatment. By W. PALIN ELDERTON, F.I.A., and SIDNEY J. PERRY, A.I.A. *Issued.*

Questions of the Day and of the Fray. Price 1*s. net* each (VII excepted).

I. **The Influence of Parental Alcoholism** on the Physique and Ability of the Offspring. A Reply to the Cambridge Economists. By KARL PEARSON, F.R.S.

II. **Mental Defect, Mal-Nutrition, and** the Teacher's Appreciation of Intelligence. A Reply to Criticisms of the Memoir on 'The Influence of Defective Physique and Unfavourable Home Environment on the Intelligence of School Children.' By DAVID HERON, D.Sc.

III. **An Attempt to correct some of the** Misstatements made by Sir VICTOR HORSLEY, F.R.S., F.R.C.S., and MARY D. STURGE, M.D., in their Criticisms of the Galton Laboratory Memoir: 'A First Study of the Influence of Parental Alcoholism,' &c. By KARL PEARSON, F.R.S.

IV. **The Fight against Tuberculosis and** the Death-rate from Phthisis. By KARL PEARSON, F.R.S.

V. **Social Problems: Their Treatment,** Past, Present and Future. By KARL PEARSON, F.R.S.

VI. **Eugenics and Public Health.** Lecture to the York Congress of the Royal Sanitary Institute. By KARL PEARSON, F.R.S.

VII. **Mendelism and the Problem of Mental** Defect. I. A Criticism of Recent American Work. By DAVID HERON, D.Sc. *Issued* (Double Number). *Price 2s. net.*

VOLUME IX OCTOBER, 1913 NOS. 3 AND 4

ANOMALIES OF PIGMENTATION AMONG NATIVES OF NYASALAND.

A CONTRIBUTION TO THE STUDY OF ALBINISM*.

BY HUGH STANNUS STANNUS, M.D. Lond., etc.

I. *Classification.*

ON page 148 of Part I of the *Monograph on Albinism* by Pearson, Nettleship and Usher from the Department of Applied Statistics, University College, London (Dulau & Co.), 1911, Professor Karl Pearson remarks, after dealing with some cases of albinism reported by myself and embodied in that work, "It is impossible to study Dr Stannus's data without being impressed with the various grades of albinism to be found in the dark races." It is particularly with the wish of drawing attention to this point that I now publish notes on cases of albinism which have come under my notice since the appearance of the above monograph. Pearson notices that there appears to be a greater variation in the albinotic characters among dark-skinned peoples, and seems inclined to attribute it to the fact that the various lesser grades of albinism are more readily distinguished among them than among fair races. With this I agree, and, as my own cases will show, intermediate forms between more or less well-marked types are found in numbers. This is a matter which I think has largely escaped the observation of those who have published records of cases. I am of opinion that for this reason consideration of the observations made by a single individual on a number of cases from a single country is more likely to give a correct impression of the subject than a mass of material drawn from many sources.

The difficulty in defining albinism lies in the fact that short of a microscopical examination of all the tissues of the body it is impossible to say whether pigment is present or not. Dealing with the tissues more especially concerned in albinism,

* This paper formed part of a thesis presented for the degree of Doctor of Medicine at the University of London. [The absence of Dr Stannus in Nyasaland has thrown upon the Editor the responsibility of editing these pages for press. Names have not always been spelt uniformly in the type-written manuscript, and at various other points the Editor has badly needed Dr Stannus' help. The Editor has ventured to add a few footnotes in square brackets on topics, which he would much preferred to have previously discussed with the author. ED.]

that is, those tissues which normally contain sufficient pigment to be evident to the unaided eye, it has been found that, even when they appear to contain none to the unaided eye, on microscopical examination a certain amount of pigment may be present; such has been demonstrated in the case of the eye and hairs.

Pearson divides albinotic negroes into the following classes:

(I) Complete albino.

(II) The spotted (freckled) albino.

(III) The blue-eyed white-skinned negro with photophobia and nystagmus but not red pupils.

(IV) The yellow-eyed white-skinned negro.

(V) Xanthous negro.

(VI) The partial albino or piebald negro.

Considerable confusion I think has arisen over the matter of the colour of the eye, which in my opinion invalidates some descriptions and also probably the above classification. "Eyes said to have the pinkish colour common to albinos," "the eyes are pink," are expressions commonly found in reading descriptions. Again, Sir H. Johnston in referring to Nyasaland writes, "cases of Albinism when the hair is yellowish white, the iris of the eye pink and the body skin an unwholesome-looking, reddish white are not common." When I say that I have collected a very large number of albinos in that country and that I have never seen a case presenting this characteristic of the iris, I think it may be assumed that in this case as in some others a lack of precise observation has led Sir H. Johnston to make an incorrect statement.

The same is, I think, true of a number of the loose statements made by other persons. They know that in typical European albinos and in, say, white mice the eye is "pink," and when they see a native of Africa with white skin and hair in whom it is difficult to get a proper view of the eye owing to nystagmus and photophobia they assume that the "eye is pink."

Dr Turner in describing cases of albinism seen in S. Africa, does not specify the colour of the irides and pupil in individual cases but refers only in general terms to the peculiar pink colour of the pupil, yet from the context it is evident he is speaking of the red reflex obtained through the pupils with the use of artificial light, though Pearson does not make this point clear.

The point I wish to emphasize is that in all cases noted by myself, when seen in ordinary daylight the pupil appears black and there is no red colour seen through the pupils under these conditions as seen in European albinos. I believe this to be true of African albinos in general, though I do not altogether doubt the occurrence in rare instances of the other phenomenon. The explanation lies, I think, in the fact that absolute albinism of the eye in negroes is very rare and that the amount of pigmentation present is sufficient to prevent a large light entry and reflexion: that is to say, the number of complete negro albinos who

will come under Pearson's Class (I) is very small, and this type is not found in Nyasaland at all events. He may on the other hand be referring to the presence of a "red reflex" obtained during ophthalmoscopical examination *.

When examined with artificial light and ophthalmoscopic mirror, all those of my cases approaching complete albinism show a "red reflex," differing from the dull red grey in the normal native in being brighter and devoid of the element of grey to a variable extent; and in some the brightness being greater than in the ordinary European, simulating that in a European albino.

The reason of the absence of light reflex (red pupil) in daylight is I believe due to the impediment to sufficient light entry caused by pigment on the posterior surface of the iris as much as to light absorption by pigment in the choroid.

In all my cases approaching the condition of complete albinism the irides are of a liquid steel blue grey colour, in some associated with a peripupillary ring of orange colour.

This I find corresponds to the descriptions given by observers in other parts of Africa whose writings bear the impress of being the result of careful observation.

I have considered these points at some length as I do not feel that the working classification given by Pearson clearly defines itself.

In Class (III) what does "not red pupils" mean? Does it mean the pupil is not red in the sense that the white rat's is, or does the expression refer to cases in which no red reflex can be obtained with artificial light?

In which class are to be placed cases which though they give this red reflex, have blue irides, in (I) or (III)?

The brightness of the red reflex varies very considerably, and I shall therefore make out one class which includes all those cases with white skin, white and yellow white hair, pale blue grey irides which give with artificial illumination a red reflex of any degree as opposed to a grey one. I shall also include in the same group as a sub-class those albinos with the above characteristics in whom, in addition, there have developed spots of pigmentation, so called Spotted or Freckled Albinos; I do this because the condition is not uncommonly found in the albinos of this country beyond the age of infancy.

I think it may be of use to make a group under Class (IV), that is, cases with yellow eyes but white skin.

* [I think Dr Stannus is incorrect at this point, the red reflex from negro albino eyes was noted in a number of cases even before the invention of the ophthalmoscope and has frequently been noted since by trained observers not using artificial light. Our albino dogs have eyes almost identical for amount of pigment with those of the albino negro, and in three cases out of four the observer thinks they have dark pupils, on the fourth occasion the red reflex is caught because the incidence of ordinary light is suitable. If four or five albino dogs are in a room together and they are then excited, any observer sees the reflex at all turns. My classification was, of course, a broad one but I think a serviceable one as it roughly graded the albino negro into two classes according to the intensity of ocular pigmentation, i.e. by manifest or non-manifest red reflex, observable by the traveller without apparatus. ED.]

Under my own Class (III) I shall consider cases of Xanthism as I understand that condition. Again I would refer to Sir H. Johnston's misleading remarks, "occasionally there are cases of positive Xanthism or a state of colouration similar in a much less degree to albinism—namely that wherein the colour of the skin and the iris of the eye is quite a light yellowish brown; this type is very much admired by the negroes." It is the last sentence which shows the incorrectness of the deduction, though this of course would not be appreciated by anyone who does not know the country.

The light coloured woman admired by natives has a yellow skin, a rather cold colour, with irides of a light brown, all the hair, however, is always black as in the average native. People of this colour are more common among certain tribes—the Achikunda, those of Senna on the Zambesi river, among the Yaos and the ruling families of Momberas Angoni—and I look upon this skin colour as a racial characteristic and not as a feature of Xanthism. Whether or not albinism is more commonly associated with such peoples I am unable to say, though Pearson thinks there is some reason to believe that blondism does exist side by side with albinism, occurring in the same districts and, it has been asserted, in the same stocks. This was first pointed out by Pritchard and has since been supported by Pearson, who quotes cases of men from Nyasaland observed by Turner in South Africa; the latter however remarks that these cases had "pubic hair brown," "in other cases the skin was a light rich red colour, wool on scalp a light yellow, the eyebrows and lashes reddish brown"; these cases I should immediately admit as cases of Xanthism but not those referred to by Sir H. Johnston, who does not mention hair colour and the skin colour of whose cases is yellow rather than reddish brown.

By Xanthism, then, I mean a condition characterised by a skin colour of reddish brown, red, or warm brown colour, scalp hair a golden or reddish brown, the same being true of the face and body hair, the irides being a light brown or hazel*.

Just as cases of imperfect albinism or partial albinism have to be recognised, so there are degrees in Xanthism, as in one of Turner's cases, but I look upon the distinctive points about the affection as the peculiar reddish "warm" colour of the skin, generally if not always associated with a similar colour of all or some of the hair; the yellow-skinned natives referred to above, though lighter in colour than many cases of true Xanthism, very seldom freckle as Xanthous natives do.

Dr Turner has also observed a number of natives with light hair "or down" on the temples, a fact which Pearson seems to think should be taken as significant as one of the first signs of a lighter pigmentation. With those downy hairs I am perfectly conversant but believe them to be merely an overgrowth of lanugo hairs. These fine downy hairs moreover are found in many natives elsewhere on the

* [The difference between Dr Stannus and myself appears to be that while I speak of a red or yellowish red skin, red, brown or even yellowish brown hair and brown to yellow irides as cases of Xanthism, he excludes anything beyond golden brown in the hair, or beyond hazel in the irides. Ed.]

body; they are always of the lanugo type and generally so small as to pass unnoticed but they will be seen to be of a lighter colour, golden to brown, and are quite distinct from the curled "woolly" hairs of the body skin.

There is another type of skin pigmentation for which a class has to be made. I refer to cases which with a light brown skin colour, not red and not yellow, and light hazel irides, have hair of a dirty yellow colour but with no red tinge in it. Some of these cases have photophobia and nystagmus pointing to deficient pigment in the eye. Typical cases form a well-marked group though I shall later show there are no absolute distinctions to be made between any of the groups or classes and that intermediate cases are met with all along the line. This last class is characterised by the fact that at birth the skin and hair are practically white but that slow darkening has occurred. The same may have been true of the irides but native observation upon which I have had to depend mostly in these cases has not been sufficiently good to remark this point.

These cases resemble the remarkable case of Dr Ascherson's referred to by Graves and quoted by Pearson, and are, I think, akin to those described by the same author as occurring in Papuans and elsewhere, in which the skin is darker than in the first form (ordinary typical albino) and in which nystagmus may or may not be present, where freckling is absent or at any rate not conspicuous, while the hair is tow coloured. Similar cases have been reported from Malay.

I have considered above albinism of various degrees but always affecting the whole of the skin, hair and eyes, that is to say cases of complete and incomplete albinism. It now remains to deal with the condition in which the affection is limited in extent, what is termed Partial Albinism. According to the extent of the albinotic characters this class is sub-divided by Pearson into Piebalds and Spotlings; the leucosis may be complete or incomplete.

It is therefore a difficult matter when considering cases to be sure under which category any particular one should be placed. Some doubt was entertained as to the existence of Piebalds until quite recently, when Seligmann's Papuan case, Gilbert Smith's English case and a family I have described from Nyasaland established the undoubted existence of Piebaldism. Persons showing a much less extensive area affected, often a single patch, have been conveniently called Spotlings and such are not uncommon in Nyasaland. It is to be noted that the area or areas affected tend to be the same in different cases. In Piebalds the belly and chest are involved and often associated with a flare on the scalp and "garter" or "stocking" markings. In Spotlings this arrangement and extent may be wanting. The various members of the Nyasaland family will be seen to present intermediate forms. This distinction into two groups is warranted by the fact that in a number of cases where the affection is inherited the type is preserved.

The resemblance of the Nyasaland Piebald family to Bishop Harman's family is worthy of note, and it has been suggested that such cases of human piebaldism

"are intensified representations of a light-bellied quadruped, that we have to deal with a case of palaeogenic inheritance," a subject associated with the name of Sir J. Hutchinson.

It will be found that I have set apart the Spotlings in which the penis only is affected as they form a special group owing to their great definition and owing to the little attention that has been paid to them.

I have pointed out, as will be seen later, that the pathognomic feature is the very light "pink" colour of the affected areas and their sharp margins.

My own working classification, therefore, is:

Complete albinos.	(I)	White skin, straw coloured hair, liquid blue iris with red reflex through pupil when examined with artificial light.
Spotted albinos.	(I*a*)	Cases as above showing spotty pigmentation of the skin.
Partial incomplete Albinism.	(II)	As above but the iris light yellow or hazel, photophobia and nystagmus not always present.
Xanthism.	(III)	Reddish brown hair and skin, light brown irides.
Incomplete Albinism.	(IV)	Skin light brown, not red; hair dirty light yellow, not brown; irides hazel.
Partial Albinism.	(V)	Piebalds.
	(VI)	Spotlings.
	(VII)	Spotling of penis.

II. *Illustrations of Anomalies of Pigmentation.*

I shall now give notes of cases and discuss the group to which each belongs. Some cases of leucoderma are then described after which the relation of leucoderma with albinism is discussed and the theories as to pathogeny of the conditions brought forward.

The cases of albinism so far recorded from Nyasaland are as follows:

A. 1. From Kayuni's, N. Nyasa, by Dr Davey (p. 145 *Monograph*).

2. Charlie* of Chikulamayembe, N. Nyasa, by Dr Davey (p. 145 *Monograph*).

3, 4, and 5. Cases mentioned by Mrs Young of Karonga, N. Nyasa (p. 145 *Monograph*).

* I have seen this case and would add to the description already given that the scalp hair of a light straw colour grows down over the forehead to the eyebrows. The whole of the back and limbs are covered with semicurved very light straw-coloured hairs. Irides light grey blue with a light orange-coloured pattern. Pupil black in daylight. Teeth irregular, ill-formed, and carious. The pedigree is given as Fig. 7 of Plate XXI, II 2 and II 4 were brothers, there are no other known albinotic members of the family.

B. Cases seen and reported by myself and included in the *Monograph on Albinism* (to which pages and Figures in Pedigree Appendix refer):

1. Girl aged $4\frac{1}{2}$ years, daughter of Dingo, Chiromo; Fig. 426.
2. Boy aged $2\frac{1}{2}$ years at Zomba; p. 148 ftn.
3. Chesiwandiri of Maumi, S. Nyasa; Fig. 430, Plate Y (77).
4. Ng'ombe, S. Nyasa; p. 147, Fig. 427, Plate Y (80).
5. Child of Mtila, S. Nyasa; Fig. 431, Plate Y (79).
6. Mbwerachaji } of Mvera, C. Angoniland; Fig. 428, Plate Z (84).
7. Kantaye }
8. Old man of Nkoma, C. Angoniland } Fig. 432.
9. Old woman " " }
10. Young woman of Nkoma, C. Angoniland; Plate Y (78).
11. Soto of Angoniland; Fig. 433.
12. Sikumbiri of Blantyre; Fig. 434.

In the families of these cases 6 other cases occurred making a total of 18.

13. The Piebald Family of Florence Bay; p. 250, Plates H (21), RR (147)—(150).

C. The following cases are now reported for the first time:

1. Alice.
2. Saiti.
3. John.
4. Pingo.
5. Chibwana.
6. Tom.
7. Ali.
8. Chikoya.
9. Nderemani.
10. Mateuta.
11. Child of Annie.
12. Moyichandi.
13. Sawali.

Case C 1. Alice of Chikanji's Village, Dwangwa, Awemba Tribe. Pedigree, Plate XXI, Fig. 11.

The father and brother, a boy aged 6 years, are of average colour, the mother is rather light-skinned, *café-au-lait.* No relatives are known to be albinotic.

The albino (Plate XVIII (25)) is a child aged 8 months, she has a pinkish tinge to a very white skin, no vestige of pigment visible. Hair as nearly white as possible, not woolly on head. Irides a slaty blue with orange pigment network round the pupils. Ophthalmoscopic examination impossible. Red reflex not present in ordinary daylight. Lateral nystagmus and photophobia marked. There are some blisters and sores on neck and ankles, the result of sunburn; on body and limbs fine white downy hairs present; a small umbilical hernia present.

This child, who presents complete albinism of the skin and hair, shows the peripupillary orange-coloured network, but no trace of brown pigment in the iris. At the present time, in the absence of any opportunity for making a microscopical examination of irides, showing this condition, I am a little doubtful how such cases should be classed. I think it is quite possible that this colour network is not due to granular pigment at all, but is caused by diffuse pigment such as is found in many specimens of red and golden hair, as will be mentioned later. This child is therefore provisionally placed in Class I.

Case C 2. *Saiti of Zomba. March,* 1910. See Pedigree, Plate XXI, Fig. 1.

The maternal grandmother is said to have been a light-skinned native (I 2).

The mother's skin colour is light brown; Bwanali (III 6) and William (III 5) are dark brown; the first four children are dead and other members of the family were said to be of average colour, a dark brown. Saiti (III 7 on Plate XXI), aged seven, is practically the same colour as the back of the hand (not sunburnt) of a fair European, there are some blisters on the shoulders and arms but no freckles; skin thickened in places. The body and limbs are covered with fine whitish yellow hairs. Hair of head a dirty straw colour; eyebrows and lashes thick and much curled. Eyes; lateral nystagmus present, day-blindness marked, irides bluish-grey, with a faint yellow pigment pattern; pupils black. With ophthalmoscopic mirror a red reflex shows through the pupil, but not through the iris. Fundus a light pink colour, possibly a little lighter than in the average European. Teeth very bad. Abdomen prominent, small umbilical hernia present.

This boy is a very characteristic case of Class I, but shows the orange colouration in iris.

The frequency with which albinos have body hair and also bad teeth will be mentioned later. The red reflex was here obtained and the fundus, though found to be even paler than in an average European, still contained enough pigment to prevent the choroidal vessels being seen.

Case C 3. *John of Mwandanas near Piupiu Hill, Zomba. January 6th,* 1911. *Mother and Father Anguru tribe.* Pedigree, Plate XXI, Fig. 5.

With the exception of the first child, which is said to have been white, all the other members of the family are known to have been averagely dark-skinned natives. An elder brother (III 3) was seen and found of average colour.

John (Plate XIX (28)), aged 6 years, a complete albino. Skin colour that of a fair-skinned European; in places harsh and dry, smooth, with cracking and some blisters from sunburn. On the lower part of the abdomen, especially on the outer surfaces in front and on adjacent surfaces of thighs, numerous warts are present, mostly small, only slightly raised, and rather flat-topped. Others are larger, as is one situated over each clavicle, there are also a few which are filiform in character. They appear brownish in colour. The body and limbs are covered

with light flaxen hairs semicurved, which, on microscopical examination, were found to contain no pigment granules. The hair on the head is very light flaxen colour, whisker areas covered with curved hairs. Eyelashes a golden brown colour, microscopically, no granular pigment in hairs*. Eyes: vision good in moderate light and pupils black, nystagmus present, red reflex present with artificial illumination. Irides, peripherally slaty-blue, round the pupil a light brown. The tongue has a bifid extremity. Teeth: two upper front incisors large, with serrated edges; all the front teeth above are spaced.

The child presents a fairly typical appearance of Class I save for the partial pigmentation of the irides, and on this account it must be considered as an intermediate between I and II. I am not sure whether the presence of the warts is of any significance; as far as I know, they have not been previously reported as specially associated with albinism. Note character of tongue and teeth.

Case C 4. Pingo of Chitalu, Liwonde, seen Zomba 18th August, 1910. Pedigree, Plate XXI, Fig. 2.

The parents, aunts and uncles of this case (II 3) and remoter members of the family are said to have been ordinarily dark-skinned natives. A cousin, a child of a maternal aunt, now dead, is stated to have been an albino. The eldest (II 1) of the family, of which Pingo is the second born, was albino, but is now dead; one other brother and three sisters normally black. Pingo is married to a normally dark native woman, who has had three children by him, all living and all dark.

Pingo (see Plate XIX (26) and (27)) is a man of 25 years of age. The skin of that part of the body covered by the loin-cloth is still very white, resembling the white skin of a European, and is soft and unaltered. The rest of the body shows changes in the skin of varying degree, the skin of legs being less affected than elsewhere; it is rough, harsh, scaly, cracked and creased with sunburns and superficial ulcerations; the skin is badly kept and dirty. There is hair all over the body and limbs, of a light straw colour; eyelashes, thick but ill-formed, are of a rather dark yellow colour.

Scattered over chest, abdomen, back, arms and neck, and also on the face are a number of areas from one third to one and a quarter inches in diameter, exhibiting a series of changes, giving the appearance of being of the nature of a pigmented atrophy; *pari passu* with the colouration, which commences as a light yellow colour, there is apparently an atrophy of skin substance, the final result being a black depressed spot, irregular in outline and with a brownish tinge at the edges, thrown into relief by the pinkish dirty white colour of the rest of the skin. Hair, a dirty straw colour, thick. Irides, a light opaque brown; lateral nystagmus present, but little photophobia. Pupils in daylight black. With artificial light, a red reflex is obtained, the fundus appearing a lightish red.

* The microscopical preparations of hairs in all cases referred to were made by Miss E. Y. Thomson, of the Biometric Laboratory, with the kind permission of Professor Karl Pearson.

When he smiles the subject has a retracted everted upper lip with fold of mucous membrane showing beneath it; the upper lip has a central linear depression instead of being pointed.

In reporting the case of Ng'ombe (B 4) it was pointed out that pigment spots hardly occurred apart from pitting and that the pitting might be the result of ulceration with scarring. In the present case, however, no history of such loss of substance could be elicited, yet a similar condition obtained. Both these cases are adults, whereas my other cases of spotted albinism are children in whom the pigmented areas are not depressed. Will these children develop depressed areas? I think they very possibly will. I believe that localised changes in the skin occur and cause the depressed areas, the pigmentation occurring before or at the same time, but whereas the pigmentation is visible in childhood the early results of depression are not seen. The irides of this man are a light brown, that is to say, more pigmented than in Ng'ombe, who only had an orange peri-pupillary ring. This case is included in Class II, though the iris colour makes it an intermediate type. A portion of skin including in its centre one of the pigmented spots was excised from the arm of this case and preserved*. Vertical sections were cut and examined microscopically stained and unstained and except for the complete absence of pigment granules the leucotic skin as a whole appeared in no way to depart from the normal. The several layers of the epidermis were not very thick, the dermal papillae and processes of epidermis between them being little developed but regular.

The structure of the skin at the site of the pigmented spot (freckle?) differed, however, very considerably. Here the dermal papillae are irregularly prolonged towards the surface so that the epidermal processes appear much longer but the rest of the epidermis thinner. The condition might be considered to resemble a wart which was "in" the skin and not "raised" above the surface.

Pigment granules were seen to occur in all the layers of the epidermis, most numerous in the cells of the downward processes of the epidermis, especially in the basal layer of these cells, which envelop the dermal papillae. They were seen to be more abundant in the peripheral parts of the cells and most densely packed at the deeper poles of the cells.

The granules were of a brown colour and where massed, appear as a homogeneous black mass.

In the subjacent dermis some cells were seen to contain pigment granules, they were few in number and the number of contained granules was small. Passing from the middle of the pigment patch where the aggregation of granules was greatest, they diminish in number as the periphery is reached. The diminution in the amount of pigment was seen to be gradual to one side of the section, the gradation taking place over several millimetres with some isolated

* See Plate XVII (21) and (22) for microphotographs of albinotic skin and "freckle."

masses of cells containing granules beyond. To the other side of the section, the transition is accomplished in 0·2 mm. This particular pigmented spot of skin was but very slightly depressed below the general skin surface, and though it is but a single observation I think the changes described would probably be found in all examples of the condition.

This is I believe the first time a "freckle" from a negro albino has been submitted to microscopical examination and it would be unwise to draw conclusions from this single observation.

Case C 5. *Chibwana of Chikala. July,* 1910. Pedigree, Plate XXI, Fig. 4.

Mother and father and two brothers and a sister all normally dark-skinned. The maternal great-grandmother had said there was an albino in the family ancestry.

Chibwana (see Plate XVIII (24)), a boy aged 4½ years (1910), has a dry cracked skin the same colour as a fair European; pigmented patches are gradually appearing on the face, also but less marked on chest and back and less still on the limbs. These patches are irregular in outline, black or dark brown in the centre, and a lighter brown at the periphery, but they are not depressed. Hair on head reaches to eyebrows, there is hair also on the cheeks, back, arms and legs, with little on chest, semicurved and white; on microscopical examination, it is found to be devoid of granular pigment. Eyelashes long and curled, of slightly yellower tint than hair on scalp. Irides, olive green with buff yellow pigment ring round periphery; pupil, not quite circular in outline, appears black in daylight. Lateral nystagmus present. On ophthalmoscopic examination a red reflex is obtained through the pupil, none through iris; fundus pink as in European.

I think in this case we see an early stage of the condition found in Pingo, which may progress until a condition similar to Ng'ombe is reached.

Case C 6. *Tom of Chifira's Village, Bandawe, W. Nyasa.* Pedigree, Plate XXI, Fig. 8.

The mother and father, former dead, were both of average colour. No history of albinism in the more remote ancestry. The first three children were albinos but were killed at birth according to custom among the Atonga. The fourth child was allowed to live owing to more civilised ideas of the chief, but died at the age of 15 years.

Tom, the last born, a boy aged 14 years (1909) is mentioned in the *Monograph* and a photograph by Dr Robertson there reproduced (Plate PP (142)), but the history of the case is not given *. The skin of non-exposed parts is a pinkish white colour, elsewhere it shows sun-tanning and also a considerable degree of spotty pigmentation, the colour being not a brown to black, as seen in other cases, but a golden yellow like ordinary freckling in a European. Irides a light hazel

* [Given as Fig. 648 in Part IV of the *Monograph*. Dr Robertson states that there were two *normal* siblings and does not mention the dead albinos. Ed.]

brown. Photophobia and lateral nystagmus present. Dull red reflex present with artificial light. Vision very fair. No hair on body. Hair very white as opposed to the usual light straw colour; microscopical examination showed it to be without pigment.

The lad, though one of the lightest haired cases I have seen, had hazel irides, though, be it noted, as nystagmus and photophobia were marked, he would be placed in Class II, but owing to his spots, he is considered in Class I *a*. The spots of pigmentation present, however, do not approach in colour those on the younger boy Chibwana. Tom lived near a mission station, and I think may have worn more coverings than the other cases, who wore only a loin-cloth. This series of five albinotic siblings, all the children born, is a remarkable example of albinism in a family.

Case C 7. *Ali of Mwaimba, near Mvera, S. Nyasa; Yao; male, aged* 16 *years.*

Three elder brothers, one younger, and one younger sister, mother and father all very dark. No more remote members of the family known to have been albinotic. The skin colour of the subject is a light "warm" *café-au-lait.* He is covered with fine golden hairs all over the body; hair of scalp, face, axillae and pubes a dirty straw colour. Irides a light reddish brown. Nystagmus present. Pupils black. Using an ophthalmoscope in a dark chamber, a red glow through pupil is obtained, but none through the iris. The optic papilla appears normal. The fundus reflects a red glow, a little lighter than that of a European; choroidal pigment is almost completely absent, allowing the choroidal vascular network to be completely visible, the small amount of pigment being collected into small deposits lying between the vessels.

The salient features of this case are a skin and eye colour which would cause me to consider it one of Xanthism, the hair also is rather dark for Classes I or II, but, on the other hand, the fundus is found to have less pigment in it than most other cases. He must therefore be considered to be an intermediate form presenting affinities with Class II and Class III. He has hairs all over the body, which I look upon as a condition very commonly associated with albinism.

This case I consider an important one in linking up the several types.

Case C 8. *Chikoya of Mulowi, W. Nyasa. November,* 1909. Pedigree, Plate XXI, Fig. 6.

Chikoya, a man of 25 years of age, is stated to have been, as a child, quite white, but has gradually darkened to present colour, of dark sunburn in a European; skin fine, not thickened nor cracked. Fine golden hairs are present on the arms, legs and shoulders; there is a light yellow golden moustache and beard; the hair of the scalp is a dark brown colour, as is also the pubic hair, but on microscopical examination it is found to contain only diffuse pigment and no granules. Hair said not to have darkened with age. Irides a light hazel; no photophobia; no nystagmus, vision good.

In this man we see a case to be placed in Class IV, a group characterised by light brown skin colour (not red), in adult life, with nearly always a history of darkening with age; the irides are hazel, the hair, though it is a dark brown colour, contains no pigment granules.

Case C 9. Nderemani of Namalwere Village, Fort Maguire.

This man, aged 25 years, was seen in Zomba in 1911 when returning from work on the South African mines.

Skin a rich "warm" brown colour; hair a rich red brown hue; axillary hair a dark brown; there are a few golden coloured hairs on the body and limbs. Irides, dark hazel; no nystagmus; vision good. There is a small umbilical hernia present. He says he has darkened with age and more rapidly since going to the mines than before.

This is a typical case of Xanthism, Class III. The hair, on microscopical examination, was found to contain few pigment granules, but a considerable amount of diffused pigment.

Case C 10. Mateuta of Chiri, W. Nyasa.

Father of average colour. Mother states that she used to be "white" when a child; at the present time her skin tint is a light brown; irides hazel; hair black.

Mateuta, a child aged 2 years; skin colour a very light brown, face of lighter tint than body; irides greenish hazel, no nystagmus; hair light brown yellow with red glint in it; eyelashes and brows of a lighter shade; no hair on body. As an infant, is said to have been nearly "white." Vision apparently good. Mother and child constitute, I think, a short series, showing a progressive pigmentation, of value to the general subject.

Their histories show them to have possessed at birth a white, or almost white skin, which has become pigmented with age; the child's skin has not yet reached the degree of pigmentation of the mother's. The child's hair is a light brownish yellow, which is becoming darker; whether it will actually become black remains to be seen, but I expect it will. The same is true of the irides. The mother's are hazel, the child's are greenish hazel, strongly suggesting that increase of pigmentation is in process of taking place.

Case C 11. The child of Annie and Annausu of Nkumbira, Nkata Bay. Pedigree, Plate XXI, Fig. 12.

All relatives dark; no other cases of albinism in family. The child (II 12) aged 8 years, is stated when born to have been almost white and to have darkened since; the same is said to be true of the hair. The face is a very light sallow earthy colour with dark freckles; skin of body a warmer tint of brown. Hair of scalp dark brown. A few fine downy hairs are seen on the body and limbs. Irides hazel; no nystagmus; vision good.

This child, aged 8 years, can, I think, be taken with Mateuta and his mother and might be placed in the series as intermediate between the mother and her child. Together they seem to illustrate simply a delay in normal pigmentation, a phenomenon to which I do not think attention has been drawn, but one which must be taken into consideration when any attempt is made to theorise on the cause of albinism.

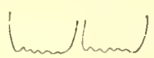
Case C 12. *Moyichandi of Sangano, Blantyre, Yao.* Pedigree, Plate XXI, Fig. 3.

Moyichandi, aged 18 years (II 5), is the brother of a case (B 12) Sikumbiri (II 6) previously reported in the *Monograph,* Fig. 434.

The skin is a light but "warm" reddish *café-au-lait* colour distinct from any normal native skin colour; irides hazel; no nystagmus. Hair light brown in colour, which on microscopical examination was found to contain fairly abundant granular pigment and a considerable amount of red brown diffuse pigment.

Sikumbiri, a year younger than his brother, has a lighter tint of skin, not so red; the irides are light hazel and nystagmus is present; hair dirty light straw colour. Both have hair on body and limbs of a yellow colour. In these two brothers we see an excellent example of the connection between two grades of deficient pigmentation; the elder I should class immediately as a case of Xanthism, whereas his younger brother must be placed as an intermediate between Classes II and IV. See Plate XVIII, Fig. 23.

Case C 13. *Sawali of Ntiya Village, Mlumbi, Zomba; Yao,* 1911. Pedigree, Plate XXI, Fig. 9.

Sawali, II 1 (for photograph see Plate XII, 1 and 2) and his younger brother, II 2, aged 13 and 5 years respectively, of whom the elder one only was seen, are said to resemble one another. Sawali seen in January 1911; hair of scalp from light yellow to brown, fairer on neck and at the edges of the hairy scalp; eyebrows and lashes black, no hair on body, pubic hair black; no albinotic characters in the skin, i.e. the skin is of a slightly lighter tint than the average native, but well within the limits of normal. Irides medium brown; no nystagmus; vision good. Teeth, all the incisors are serrated thus:  ; they have the appearance of having ill-developed enamel.

Seen again two months later. He had in the meantime cut his hair very short, almost shaved, and allowed it then to grow again; the soft less-curled hair on the temple is a mouse-brown colour, the newly grown, tightly curled wool of the scalp is of a yellow colour, becoming brown near the roots; scalp colour is the same as the skin. He states that he was born with a white skin which has darkened since.

Individual hairs were found to vary in colour; some very light golden hairs were found microscopically to contain a yellow diffused pigment with very fine

granules in fairly large numbers; brown hairs contained brown diffused pigment and abundant granules; some light hairs contained comparatively few granules, fine and widely scattered. There are some spots of pigment-like freckles of dead black colour, a few on neck, chest and back. On ophthalmoscopical examination a slight dark grey red reflex is obtained through the pupil; the fundus is a dull red, rather lighter than in average native.

Cases of partial albinism in which the hair only is involved but completely must be of rare occurrence, and I know of no case among African natives; such a striking figure, if seen, would surely have been recorded.

Note should be taken of the character of the teeth and the presence of body hair. This case should be considered together with the last cases above mentioned.

III. *Piebalds and Spotlings.*

With the exception of the Florence Bay family referred to earlier in this paper no other cases of Piebaldism have been seen.

The following cases are examples of spotlings that I have come across in Nyasaland. In the first group mention is made of white tufts of hair: in the second, spotlings of the skin are described.

(A) Several cases of natives presenting congenital white tufts of hair have been seen; they resemble each other almost exactly in that the canities occurred in the form of a median-placed flare on the anterior half of the scalp more or less triangular in arrangement, base in front. There were generally also a few white hairs scattered over the rest of the topmost portion of the scalp.

One such case also presented right and left helical fistulae, a congenital anomaly of no significance in this country, as it occurs in about 4% of the population.

Salule, of Chitanji Village, Chiradzula, showed two small tufts of white hairs, one placed in the median line, the other a little to the right and behind it, at the summit of the scalp; the scalp beneath them was a little lighter in colour than the surrounding. This man states that he was born like this and that one of his brothers, but no other member of the family, is similarly affected. The white hairs submitted to microscopical examination were found to be devoid of pigment.

G., the wife of a private in the King's African Rifles, aged 25, was seen to have white hairs scattered over the scalp which were said to have been present since birth; she presented no other albinotic characters.

Chitesima was another example of this condition. Microscopical examination showed the white hairs to be devoid of pigment, granular or otherwise.

H., a girl, aged 6 years (1909), from a village on the Lake shore in West Nyasa, presented a curious and interesting condition exemplifying partially

deficient pigmentation. The affected areas were the right side of the upper part of the face, two small patches on the same side of the neck and a patchy condition of the fingers. (See Plate XIII, Fig. 5.)

The area on the face was sharply limited along the mid-line of the nose and forehead. The two patches on the neck were at the anterior margin of the trapezius muscle over the mid-point of the clavicle. The mottled condition of the fingers rather resembled that seen in Ziemann's leucoderma.

The general skin colour was that of cake chocolate. The affected areas were the colour of bruised chocolate. In addition to the affection of the skin, a segment of the iris about one-sixth of the total iris in extent, with the base directed towards the right side, was of a light hazel colour, the remaining five-sixths being a dark brown, the same as the iris of the other side. The hairs of the scalp, eyebrows and eyelashes were normally black. The condition was noticed at birth, according to the mother's statement, and had not altered. The mother, father and brother of the subject were all normal, and no members of the family were known to have presented any pigmentary anomaly.

I., female, aged 5 (1909). Seen at Kondowe; she presented congenital anomalies of pigmentation in that there was absence of full colouration of the skin; the back of the right hand, with the inner three fingers, the palm, and the back of the right arm, the abdomen, groin, thigh and leg of the same side were the seat of patchy deficiencies of colour, with spots quite white or pinkish white to various shades of brown, resembling somewhat the condition seen in leprosy, or the non-leprous leucodermic affection of the skin of the hands described by Ziemann. The condition was stated to be congenital and not to have altered. No other members of the family were affected.

Daus, of Chimwai village, Dwangwa River, male, aged 11 years (1909). An area of skin on the cheek, as indicated in Fig. 6, Plate XIII, is of pinkish white colour resembling that of an albino; the area is smooth and on the same plane as the surrounding dark skin. The demarcation is abrupt, with no increase in pigmentation of the skin in the neighbourhood. Over the albinotic patch and also more abundantly on the surrounding skin are distributed fine white short curved hairs, no other hair being present on the cheeks. A few white hairs are seen on the same side of the upper lip, and the middle portion of the eyelashes on the upper lid of the same side are white.

There is an area of skin just in front of the tragus of a lighter hue than the rest of the skin of the face, which is a rather dark brown black colour. Irides dark brown. History states that these anomalies of pigmentation were present at birth. They have enlarged with growth, but there has been no relative increase in size. No other members of the family known to be albinotic.

The girl "H." presents a condition of extraordinary interest. In the first place, I can put on record a case of congenital incomplete spotling albinism, that

is to say, a case with small areas of skin, congenitally deficient in pigmentation, but not completely albinotic. Secondly, the case demonstrates, I believe, a previously unrecorded condition, a partial and incomplete albinism of the iris; it is noteworthy also that the condition only affects one side, in fact the area on the face is strictly limited to the one side by the mid-line of the nose. This case will be mentioned again later.

In speaking of the other two cases I am not quite so sure of my ground.

The girl "I." presents no definite pink area of skin, but a condition resembling very much that described as Ziemann's leucoderma. It is said to have been present at birth, and the similar condition of fingers in the girl "H." supports the probability of this statement being true. The history that the condition is not changing is against leucoderma and therefore, on the whole, I am inclined to include it as a case of spotting.

(B) The special group of partial albinos characterised by an affection of the glans penis must now be considered.

My attention having been drawn to the point, the condition was then sought for among a number of adult men aged from 20—50, and including circumcised and uncircumcised, those with a partly retracted prepuce and those with a long covering. More or less complete circumcision is practised among the Mohammedan Yaos, but not among other tribes.

At the outset one must recognise that, just as in Europeans so in natives of Africa, there is very considerable variation in colour of the glans penis and inner surface of the prepuce, colour conditions which are fairly comparable to those of the mucous membranes. Just as in Europeans, variations from a delicate pink to the colour of uncooked lean bacon, from almost white to a purple red, occur; a similar variability is seen in uncircumcised natives, but always in the darker hues, that is, the skin in this situation always contains some pigment.

Dr Stachan, quoted by Pearson, states that the glans penis beneath a phimosed prepuce is always red; I am not quite sure whether he wishes to infer that the skin is devoid of pigment or not, but certainly the latter condition is the truth.

In the circumcised, the glans is practically always of the same hue as the rest of the penis and scrotum, i.e. a darker tint than the general colouring of the body.

In the uncircumcised, as a rule, according to the extent of the natural uncovering of the glans so is the degree of pigmentation.

The portion of glans normally uncovered by a naturally partially retracted prepuce is often darkly pigmented, shading off to red as the corona is reached; sometimes the general colour is a dark lean bacon-red, it may be with a purplish hue; in others, the general colour of the glans is an opaque leaden white. In

all these cases, however, all variations in colouration shade into one another, and it is evident pigment is present.

Those cases which I have considered as spotlings show areas of various extent, always of a "pink" colour, with sharply defined margins, this being the test of a departure from the normal, which I consider includes the variations in colour above mentioned.

The following is an analysis of nearly 400 adult males, between the ages of 20—50 years, being prisoners and soldiers, and therefore they can be considered as taken at random from various tribes.

379 *Adult males examined with reference to colour of glans penis and prepuce.*

(A) 122 Circumcised: All have completely black glans penis except two (*a* and *b*).

(*a*) Some spotty pinkish colour present at the seat of the operation cicatrix. (N.B. the circumcision was performed when subject was adult.)

(*b*) Chambveka, the general colour of the glans is a dark brown, but there is a light brown spot on the dorsum near the meatus.

(B) 259 Uncircumcised:

(1) Inner surface of prepuce and the whole of the glans penis dark brown or black . . . . 45.

(2) There is a gradation of colour from the corona to the meatus and corresponding part of the inner surface of the prepuce from blue red, leaden or yellowish white to dark brown or black . . . . 100.

(3) Glans and inner surface of prepuce of a uniform colour, from a leaden white to a bacon-red . . . . 75.

(4) Glans and inner surface of prepuce of about the same colour as in a fair European, but at the free edge of the prepuce shading off into dark brown or black of the general skin colour . . . . . . 20.

(5) As in (4), except that the lips of the meatus were dark brown 3.

In another case there was some deeper pigmentation about the frenum shading off into the surrounding colour; the skin colour was a dirty yellow brown, rather like that of a half-caste, with some freckling; hair black; irides dark brown . . . . 1.

(6) Spotlings . . . . 15.

From the above it will at once be seen that among the circumcised one may say no spotlings were found; case (*a*) had some alteration in the neighbourhood of a scar and is of no import; (*b*) had a condition which may be called leucoderma, or may be a congenital condition, and be called albinotic, but no history was forthcoming.

Colour variations among the uncircumcised, (1—5) came within the limits of normal variation as above described. The single case mentioned in (5) I look upon as really one of the intermediate cases towards albinism, a view strongly supported by the freckling.

To deal with cases under (6) in detail:

(*a*) Lijuni, aged 30, glans and inner surface of prepuce a very bright pink (considerably lighter in colour than in an uncircumcised European), bounded by a sharp line of demarcation at the junction of the inner and outer surface of the prepuce. There are a few white hairs in the mid-line of the beard and at the outer corners of the moustaches, but no other albinotic characters. See Plate XIV, Fig. 8.

(*b*) Erya, an Angoni, as in (*a*) but with an area of dark brown, involving the lips of the meatus, though covered by prepuce, thus differing from cases mentioned above (5).

(*c*) Matesi, an Atonga (see Plate XV, Fig. 11).

(*d*) Jim, ground colour of preputial-glans-skin of a reddish grey, albinotic area shown in Plate XV, Fig. 12. Skin colour a dark chocolate; irides dark brown; on the front of the abdomen there is a spotty condition, light and dark brown; there are a few freckles, one each on the right side of the abdomen, left chest and back.

(*e*) Chizizu, Yao (see Plate XV, Fig. 13), nothing else to remark.

(*f*) Lief. Dark brown skin and irides (see Plate XV, Fig. 14).

(*g*) Chiwia. Besides the larger patches of albinotic skin on the glans, there is a spotty pigmentary change about the frenum (see Plate XVI, Fig. 15). The skin and irides were a dark brown colour, and there were some scattered grey hairs on the scalp. Man, aged 30 years.

(*h*) Chivia, Angoni, skin and irides dark brown (see Plate XVI, Fig. 16).

(*i*) Naamon (see Plate XVI, Fig. 17).

(*j*) Yohani. There is a patchy light and dark brown pigmentary change in the neighbourhood of the frenum.

(*k*) Magombo. Red pink ground colour of glans, etc. with leaden blue spot, as shown in Plate XVI, Fig. 18, skin colour a dark chocolate.

(*l*) Asani like (*k*), only spot is brown in colour; scrotum and penis black, skin colour light brown with large black freckles, one on each leg, and some on body and face.

(*m*) Mwenengwi. Tightly phimosed, but some spotty pink areas can be seen just on the inner surface of the foreskin.

A further case has since been seen. Binali, 1st K. A. Rifles, Zomba, 1912 (see Plate XIV, Fig. 9).

Lijuni [(*a*) above] was circumcised by me on 5th April 1911. Fourteen months later he was again seen. Examination revealed a circular patch of pigmentation dark brown in colour involving the lips of the meatus and a small area round them, three-eighths of an inch in circumference; the remainder of the glans was of a light pink colour, totally devoid of pigment as at the time of circumcision. The prepuce was carefully preserved and transverse sections prepared and mounted unstained and stained with haematoxylin and eosin and by Van Gieson's method for microscopical examination. I may say at once that, with the exception of the deficiency of pigment granules, the whole structure of the albinotic areas of the skin appears normal. In the normally pigmented part of the prepuce the basal cells of the Malpighian layer are the most deeply pigmented, some of them being so full of dark brown granules as to appear almost black and homogeneous; this is most marked in the cells at the end of the deep processes of epithelium dipping down into the dermis and also in the cells lining the lower parts of the sides of these processes. Those cells less deeply placed are less pigmented, but all the cells, even the most superficial ones, of the stratum corneum contain some granules. The granules tend to be arranged round the periphery of the cells, especially at the distal end of the cells, giving a very characteristic appearance. In the subjacent dermis a few cells containing pigment granules are seen, small in number and not very heavily pigmented. The transition from the normal pigmentation to skin absolutely devoid of pigment is seen to occur in from 1 to 2 millimetres, that is to say, it is not absolutely sudden but the gradation is rapid. The diminution in the number of granules is practically proportional in all layers of the epidermis; in the dermis the loss appears possibly to be absolute before the same change is complete in the superjacent epidermis, and the cells lining the sides of the deep epidermal processes retain a greater number of granules than the cells at the more superficial and deeper parts: see Plate XVII (19) and (20).

Prolonged beyond the last cells containing granular pigment a band of yellowish brown diffuse colour is visible affecting a layer of cells two or three cells thick at the junction of the stratum corneum and stratum granulosum.

IV. *Notes on Anomalies of Hair Pigmentation.*

Samples of hair were examined with results as follows:

Simbeleta and *Nederi,* a type of rather coarse hair, which is not very curly, recognised by the natives as turning white early in life. Microscopically white hairs were devoid of all pigment, a few brown hairs showed brown diffuse pigment, with densely packed granules, the black hairs being normally black and completely packed with black granules.

Mwana Ngala. Aged 2½ years, stated to have had yaws, and in consequence of this illness the hair at the sides of the head and temples is brownish in

colour; microscopically there was a diffuse brown pigmentation and granular pigment was fairly abundant.

Another child showed the same changes under similar conditions.

Old female inmate of the Central Prison of an age when grey hairs have usually appeared; she was noticed, however, to have some brownish hairs, and others of a bright golden colour; the black and brown hairs were found to be tightly packed with black granular pigment—the golden hairs contained light yellow diffused pigment, but no granules.

Kause. An old female lunatic, presented the same hair characteristics, macroscopic and microscopic.

Three men were noticed with a certain number of red hairs in the moustache, comparable to the hairs in the moustache of a "red haired" European.

Yotam. Scalp hair black; skin a lightish brown, is said to have darkened with age; irides a dark hazel. Microscopically the red hairs contained a certain number of fine pigment granules, but also bright red diffuse pigment, black hairs absolute black, with granular pigment; some of the hairs were partially red and partially black, each part showing the corresponding condition mentioned above.

Pamandhe. Brown moustache hairs showed diffuse brown pigment and fairly plentiful black granules; black hairs and some particoloured hairs as in case above.

Chamveka. Light brown hairs were devoid of pigment granules, there were some particoloured hairs, which, with black hairs, answered to the description above.

Bandsman, 1st K. A. R. Hairs from dermoid cyst of forehead. Hairs spindle-shaped, tapering at each end; brown diffuse and granular pigment present, but not equal to full pigmentation.

V. *Case of Leucoderma.*

I now describe a case which clinically is one of leucoderma. It is the only case showing leucodermic areas of any extent that I have met with, from which I think I may infer with reason that leucoderma is rare among these natives save in its minor manifestations.

This case corresponds very well with a text-book description of the disease—a symmetrical and progressive condition associated with hyperpigmentation. Points of interest are that the patient is sure the hyperpigmentation has occurred on sites previously depigmented. The hairs on affected parts of the skin which are naturally hairy are black. Black hairs are also seen to occur on leucotic areas on non-hairy portions of the skin.

Ouda, of Chechiwere's, male, aet. 15 years, Zomba, November 14th, 1910. Disease started 3 years ago; he was seen at that time by my native Hospital Assistant who vouches for this date. He has never had yaws. The disease

commenced as a very small patch on the right shoulder which has never spread. Now spots are seen on the body, face and limbs (Plate XII, Fig. 3).

He states that a patch appears very quickly, in a few days reaching its maximum size and that the extremities are the last to become involved. At the present time the middle two-thirds of the red margin of the lower lip are affected, there is a patch on the right side of the filtrum of the upper lip, some small areas over the left shoulder, three patches the size of a half-penny over the upper part of sternum, a few points over the chest and abdomen in front and on back; the arms from the elbows downwards are the seat of large patches; the wrists, thumbs, dorsum and palm of hands and fingers also show patchy depigmentation; patches on the left knee, the dorsum of the feet and toes, and over the ankles are also present.

These affected areas are of a pink colour as of a fair European skin; they are in some places a little crusted.

There is a lack of signs of slow depigmentation, that is, areas of all shades of dark brown to pink are absent. In association with some pink areas there is a surrounding hyperpigmentation; these the boy states were previously depigmented and pink. This phenomenon was well seen on the arms, the leucotic areas presenting black hairs as on unaffected skin. Affected areas over the manubrium sterni, though there is no hair on the surrounding normal skin, present hairs which are black. No one else in the family affected.

Under the name of "Melung" (Beta), Ziemann has described in the *Archiv für Dermatologie und Syphilis*, Bd. LXXIV. S. 163—170, Wien und Leipzig, 1905, a skin disease among negroes on the West Coast of Africa "characterised by the appearance of round, oval or irregular shaped patches of a bright yellow red colour on the skin of the extensor and flexor surfaces of the hands and feet." Here in Nyasaland I have seen a number of cases which correspond more or less with the descriptions given by Ziemann, but there are several points of difference, and I am inclined to think that that observer has made too hard and fast limits for the type of leucoderma which he has described. He says, "The predominant colour of the affected areas is of a whitish tinge and the general aspect may be compared to that of extensive superficial burns which have healed without the formation of granulation tissue. The disease begins at the age of from 10—15 years; its course is symmetrical, either attacking the hands first and then the feet or the feet first and the hands afterwards, or both at the same time. The progress of the disease is extremely chronic." He further notes that never more than four-fifths of the surface of the hands or feet are affected with pigment atrophy. The flexor surfaces are more extensively involved than the extensor, the pigment atrophy is practically confined to the hands and feet, only further slight extensions occur after 10—20 years and there is gradual loss of pigment in the hairs situated in affected regions. Sensibility to touch, temperature, and pain and muscle sense are unimpaired; the sweat glands are normal;

the electrical reactions of the affected limbs are normal and there is no evidence of disease of the nervous, blood or excretory systems.

As far as I am aware no mention has been made of the disease occurring elsewhere than among West African negroes, I therefore give some details of a few cases met with by myself.

Chisfufu. Male, aged about 36. The headman of a village near Zomba. Subject states that the disease commenced on the left side of the chest when he was 6 years of age, then both hands and arms were affected followed immediately by the legs, and it has been slowly progressive till the present time. When seen there was a spotty hyperpigmentation on the front of the chest (see Plate XX, Fig. 29) only two areas being actually of a pink hue, the same condition is seen on the back; the changes on the hands and feet are typical. The depigmentation seems to have occurred in circular areas so that on fusion the margin assumes a scalloped form.

X. Female, aged 50. States that her affection only commenced 2 years ago. On examination the left hand and wrist are found to present the typical features of the disease, and though of only 2 years' duration it is as marked as in cases of 20 years' standing. On the palm of the right hand there are some spots $\frac{1}{3}''$ in diameter of slight hyperpigmentation, signs of the commencement of the disease in this extremity.

Binti Masua. Female, aged 30, of Mponda, wife of Mwini Dowa, private 1st King's African Rifles.

She states that she had yaws when aged 8 and that the leucoderma appeared at that time, and has remained of about the same extent since childhood. On examination there are seen some black areas on the arms, apparently the hyperpigmented scars of old yaws. She has a Dupuytren's contracture affecting both little fingers, the first interphalangeal joints of both hands are thickened, and the fingers show a slight degree of ulnar deflection. The affection is less advanced in the left hand than the right. In the former the depigmentation only affects the palm, while on the right side, the dorsum and fingers are also involved. Hyperpigmentation is present as a band round the wrist, along either border of each hand across the bases of the phalanges and on the back of the first fingers and sides of the middle fingers.

Licheta. Male, aged 50, inmate of the Central Prison, Zomba, an Angoni of Dedza.

The disease is said to have commenced at the age of 17 on the palmar surfaces of the wrists, since when it has very slowly spread till the present time.

The affected areas include both surfaces of the hands and wrists, and extends two inches on to the forearm; the depigmentation is patchy in character; where the process has been most intense the skin is of a yellowish pink colour. Circular areas are seen showing all grades of depigmentation as evidenced by the different

shades of brown present. Round the areas in which the change has taken place there is hyperpigmentation more or less marked, in some areas the skin being absolutely black. There is a worm-eaten appearance on the palms of the hands which the man states is the result of yaws in boyhood. There is also compound ganglion affecting both wrists. The photograph (Plate XIV, Fig. 10) illustrates rather poorly these features.

Mundana. Male, aged 40, of Chikumbo, Mlanje, Makua.

Patient states that the affection of his hands and feet was present at birth, and has not altered since. This I think is very doubtful and probably one may infer that the disease began in very early childhood, and has not progressed of late years. He says he has not suffered from yaws. Both hands and feet are affected, the former to a small degree and the feet still less, so that it is hardly noticeable in the latter.

There are only two or three small areas actually pink, for the most part the colouration is patchy and of various shades of brown to orange. The process has involved the hands and fingers, the wristband area and the dorsum of the foot and toes on both sides.

There is a hyperkeratosis of the palms of the hands and of the skin of the extensor surfaces of the interphalangeal joints. The palms of the hands also present a worm-eaten appearance with the formation of little pits. There is likewise heaping-up of epidermis with cracking at the flexures on the palms. The hairs on the extensor surfaces of the fingers have fallen out, leaving little pits.

Y. Male, aged 40. A Swahili from Zanzibar, Range Capitao, Zomba. The hands only are involved, but in a perfectly typical way. The disease is stated to have commenced when the subject was a small boy.

Other cases have been seen but of them no notes have been kept. In discussing the relation of these cases to ordinary leucoderma it must be remembered that natives are seen with a similar patchy speckled depigmentation process affecting the skin of other parts than the hands and feet. Thus I have seen a condition along the middle line of the back or in the middle of the chest on several occasions. In the case of a girl I., aged 10 years, mentioned above, it was associated with similar speckled condition of the fingers, and was said to have been present at birth. Chisfufu, as will be seen by a glance at the sketch on Plate XX, not only had a typical affection of the hands and feet but the process also involved the trunk, all four limbs and the lips.

While it is true that there is a large number of cases of an "imperfect" leucoderma corresponding more or less to the description given by Ziemann, I find that there are many others in which the process is wider in its distribution and which are therefore connecting links between the type described by Ziemann and ordinary leucoderma. Even in such cases of long standing the areas involved are of small extent, and it is but seldom that one would see a portion of skin of

any size reduced to the pink colour such as is seen in ordinary cases of leucoderma i n natives.

It is however difficult to see where to draw limits around any particular type, though I must agree with Ziemann in setting apart the hand and foot type as a fairly definite variety. I agree with him in thinking that the disease does often start during youth, but I am not convinced that it always does so. Again it is nearly always symmetrical after it has been existent for some time, but not always so at the beginning. I believe that in most cases the hands are affected before the feet, and though in the majority of cases the progress is very slow, yet in some it may run a more rapid course for a time, and then become stationary.

I have certainly never seen more than four-fifths of the hands or feet affected. The flexor surfaces are more extensively involved than the extensor, but the change is by no means so commonly limited to the hands and feet as Ziemann makes out. I would say that it is more usual for some two or three inches of the fore-arm to be involved as well. I should be inclined to lay more stress on the character of the colour changes. Practically in all my cases there was hyper-pigmentation to be seen in areas of skin adjoining those in which loss of colour had taken place. This was generally most marked as a black band around the wrist, but was seen elsewhere.

The loss of pigment as above mentioned would appear to take place at a number of points more or less at the same or at different times, so that round areas are produced showing varying degrees of loss of colour which may coalesce with the formation of larger areas with scalloped margins. Often in the middle of such a depigmented area, one that may have been reduced to whitish-pink or yellowish orange colour, small spots of more deeply pigmented skin may be seen to remain.

Ziemann has compared the appearance to the result of extensive superficial burns which have healed without the formation of granulation tissue. Such a description I do not consider good, as under those circumstances the skin is essentially pink and never of the orange yellow or yellow white colour characteristic of the affection under consideration. Again, the laying down of pigment after burns occurs at points partly separated from one another, which appear very dark indeed, and do not resemble the darkly pigmented patches seen in this disease.

Attention might again be drawn to the fact that the history of yaws was obtained in several cases and a few show evidence of lesions almost certainly the result of framboesia, I refer to the pitted appearance of the hand and hyper-keratosis. The loss of the hair from the digits in one case should also be noted as possible evidence of a trophic element in the etiology of the affection.

Below are short notes of cases of melanodermia of congenital origin as far as their histories could be elicited.

Female, aged 17, Zomba, 1910. An area of absolute black skin is seen on the right side of the neck extending over the clavicle. No other affection of pigmentation noted.

Bimba, male, adult, Zomba, 1910. Patch of black skin on the dorsum of the left foot, over the heads of the first and second metatarsal bones.

Talo, adult, male, 1st K. A. Rifles, Zomba, 1911. A black patch of skin under the right nipple.

Mainana, adult, male. Below right axilla there is an area of skin 4—5 inches square in extent of a dead black colour.

Conclusions.

The cases I have cited go to illustrate the following facts:

1. Albinism is common among the natives of Nyasaland.

2. It is a family affection and often directly hereditary.

3. Varying grades of albinism are met with in the same family.

4. Almost every degree of albinism is met with.

5. The degree may vary with age.

6. A secondary laying down of pigment in the albinotic skin in spots is a not uncommon feature.

7. Albinism is very generally associated with an unusual development of the lanugo hairs and less commonly with mal-development of the teeth.

8. Piebalds are comparatively rare.

9. Spotlings are less rare.

10. There is a very definite class in which the spotling condition affects the penis only.

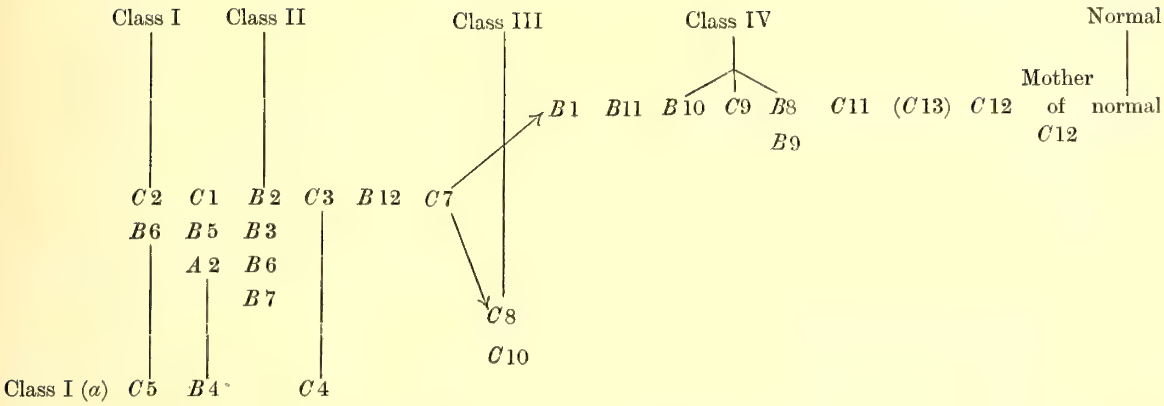
11. Widely distributed leucoderma is uncommon.

12. A localised form resembling that described by Ziemann is common, but intermediate cases between this form and a more generalised form are seen.

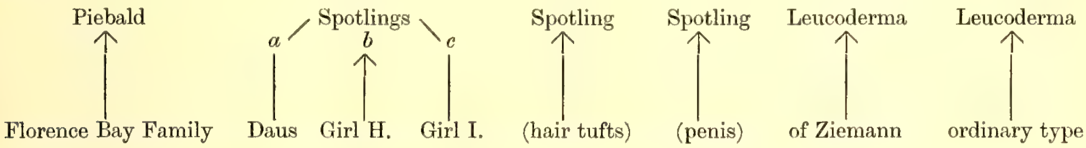
13. Some congenital cases of spotlings resemble exactly some cases of leucoderma.

The grades of albinism in their relation to one another may be diagrammatically represented; this arrangement has however no pretence to mathematical precision.

In the same way the relations of the various conditions characterised by partial albinotic characters may be roughly shown.



Partial Albinism



Coming to a more general discussion of the subject, among facts which have to be taken into consideration the following mentioned by Pearson are of great importance. Leucoderma may occur very early in life and become stationary. Pearson says "the possibility in Leucoderma of very early development followed by a stationary condition, and the fact that babies only gradually develop skin pigment so that congenital unpigmented areas may not be at once noticed, render it not so easy to distinguish partial albinism from stationary leucoderma in a European." Concentration of pigment in the neighbourhood of leucotic patches has been laid stress on as a feature of leucoderma, but it is not invariable. Symmetry and progressiveness and the combination of excess and deficiency are held by Crocker to be the characteristics of leucoderma, but as Pearson says "there can be little doubt that...leucoderma ceases to be progressive in some cases and that in such cases...the increased pigmentation at the borders is wanting." Further the symmetry "may be by no means marked" and the distribution of patches in undoubtedly congenital cases of leucosis may show a considerable degree of symmetry. Evidence of recovery of pigmentation in leucodermic patches is not wanting.

Leucoderma has further been shown to be inherited, and a case due to Dr Joseph Jones has been published in which inherited leucoderma seems "to have culminated in albinism." There thus appears to be no criterion by which any case can be judged to be one of leucoderma or of partial albinism, and the history may yield the only facts of value. For these reasons, any hypothesis to explain albinism must explain at the same time the facts about leucoderma as they are known.

In dealing with the pathogeny of the condition I follow largely Pearson in his presentation of various theories. The normal formation of skin-pigment is a question around which much conflicting opinion has grown. The site of the formation of the pigment has been variously supposed to be in cells belonging to any or all of the primary layers, endo-, meso- and ecto-derm.

Special cells have been described as acting as carriers of pigment from one layer to another and called in consequence chromatophores (Kölliker, Simon, Kerbert, Ehrmann); Ehrmann later considered these cells to be actually the producers of the pigment, and calls them chromatoblasts or melanoblasts. Meirowski's summing up of Ehrmann's views quoted by Pearson is as follows:—

(1) The production of pigment occurs in special cells, the melanoblasts, which are not identical with mesodermal or epidermal cells.

(2) The melanoblasts are products of the middle germ layer, which in part further develop there, grow into the epidermis, and there have an independent cell existence.

(3) The material which is converted into melanotic pigment comes from the blood and is haemoglobin. Haemoglobin is converted into melanotic pigment by the vital processes of the melanoblasts.

(4) Extra-cellular creation of melanotic pigment is not yet demonstrated. True melanotic pigment only occurs in an extra-cellular form by the break-up of pigmented cells.

(5) The transfer of pigment occurs by aid of protoplasmic flow along protoplasmic threads, which connect the melanoblasts with the epithelial cells. For this reason this theory is not an "Einschleppungstheorie" but an "Einströmungstheorie."

(6) Pigment is, at least shortly after its production, a body dissolved in a highly fluid colourless substance.

Unna believes that the pigment granules formed according to Ehrmann's hypothesis pass by way of the lymph stream to the epithelium.

Cohn holds that free pigment certainly exists outside the cells.

Schwalbe believes that epidermal and dermal cells can each produce granular pigment as a result of their individual activity, drawing upon the same source, though independently; he believes that: (i) a pigmentation of the cutis can occur with completely unpigmented epidermis and hair, and (ii) a pure epithelial pigmentation can occur without trace of pigment in the underlying mesoderm. He rejects the chromatophore theory and considers these cells to be transformed epithelial cells.

Kaposi also criticises Ehrmann's views to show that a haemoglobin origin for melanin is untenable. More recently other observers (Meirowsky, etc.) have come to the conclusion that epidermis and cutis pigment can arise independently, and

the melanoblasts of one and the other are independent structures. The nature of melanin has been investigated by a few observers. Hopkins and Cole have isolated a substance as a product of tryptic digestion of proteid which may be the mother substance of the proteinochromes skatol, pyridine and tyrosine. Tyrosine by the action of a ferment can be changed into a substance resembling melanin.

This ferment tyrosinase has lately been found in a variety of species, plants and animals, including newborn rabbits, rats and guinea-pigs.

The above notes on melanin are gleaned from Pearson's observations on Von Furth's summary.

Dyson has recently shown, by means of the bichromate-haematoxylin method of Lorrain Smith and Mair, that pigment is always formed as part of a complex granule which contains lipoid substance. These granules are present constantly in the basal epithelial cells, and are formed there under the influence of the nucleus as a part of the normal metabolism of the cell. The pigment is in origin a lipochrome, the melanin being the chromatic proteid portion after separation from the complex lipoid granule. He found no evidence of pigment formation in the cutis, but showed sections which illustrated the passage of lipoid substances from the epidermis into the cutis along lymphatic channels.

It appears therefore that our knowledge on melanin and its production so far points to the following conclusions:

(1) Melanin is not a direct (and probably not an indirect) product of haemoglobin.

(2) It is the product of splitting of a complex lipoid substance possibly as the result of a ferment action, akin to the formation of pigment from tyrosin by tyrosinase.

(3) The complex granules containing the lipoid substance are produced as the result of the metabolic activity of the individual cells of the epiderm.

(4) The presence of melanin is responsible for skin pigmentation.

(5) Melanin in the cutis is probably derived from the epidermis.

This is as far as knowledge goes at present with regard to melanin. How can we explain its more or less complete absence in albinism, etc.? Pearson sums up "when we consider the relative rareness of complete albinism, of the spotted or splashed condition and of Xanthism, their relatively frequent coincidence in the same stock suggests that these abnormal pigment conditions are not wholly independent, and that as a working hypothesis it is reasonable to suppose that complete albinism, partial albinism, incomplete albinism and Xanthism, all static forms of leucosis, are phases of the same process and are probably linked with leucoderma and possibly other forms of dynamic leucosis. By 'linked' we suggest that they mark the complete, incomplete, local or progressive failure of the same metabolic process, which may never start at all, never start in certain areas, or be imperfectly started, and again being started may fail to maintain itself; further

that every variety of this failure may individually or collectively be associated with certain stocks, which may either show hereditary failure of one phase, of several, or exceptionally of all phases of pigment metabolism."

This hypothesis gives a masterly summing up of the position, warranted, I think, if only in the light of my own series of cases.

If we seek the cause in some inherited defect we might suppose that some pigment-controlling determinant by its absence causes failure of natural pigmentation. Its perfect or imperfect impotence may be a racial feature of certain forms of parental albinism. This impotence may be a racial feature of certain parental stocks or produced by some degenerating influence of the somatic on the germ cells of the individual parent such as must occur in the case of an inherited character arising in the parent as a "sport."

This absence of certain determinants does not carry us very far, for the question suggests itself: how do such determinants act on the normal body metabolism? It has been suggested that some ferment necessary for pigment formation such as tyrosinase is absent.

As Pearson says, "it must however be remembered, (i) that the absence of the ferment in the albino must correspond to some abnormality in the zygote and ultimately to some defect in one or both gametes. This defect can hardly be an absence of ferment. (ii) The materials for development are supplied in utero by the mother, who may be (*a*) herself pigmented, and (*b*) produce at the same time pigmented and unpigmented twins. It would appear that she is thus able to provide the ferment for one and not for the other offspring."

One might postulate ferment determinants. Pearson suggests "that the ultimate difference between the normally pigmented individual and the albino, will be found after all to be one of structure; the failure of the normal metabolic process is due to differential structure in the albino. If there be local absence of ferment it may possibly be that the structure does not permit of its reaching its destination," and mentions in support of this view that albinos may have a considerable amount of diffused pigment in their hair.

He also says that it would be difficult otherwise to understand how in cases of partial albinism the ferment should be confined to some portions of the skin and be wanting in others. Facts gleaned from my own observations which seem to have a direct bearing on some of these points I may here refer to. I should like first to mention some observations on the natural rate of pigmentation in African natives.

Pearson states that a black baby at birth almost resembles a white one in colour, and quotes Dr Strachan*. With all deference to Dr Strachan, I most

* [This does not seem adequately to represent the views of the authors of the *Monograph*, or of Dr Strachan. The reader is referred to pp. 120—124 of that work, where much of the literature on the subject is cited. The point involved is not the impossibility of distinguishing between a just born white and just born black baby, but to the fact that the black pigmentation except on the genitals and some other parts is relatively slightly developed at birth. Ed.]

certainly must differ from him and agree with Dr Pooley, also quoted, when he says that "in newborn babies (negroes) the skin being thinner, the pink of the blood vessels is much more clearly seen, but the skin has an ashy grey colour all the same, the black is there all the time"—a phenomenon which is clearly seen in the still-born and premature births. I do not hesitate to say that a portion of skin of the new-born native of this country could be immediately distinguished from a portion derived from a new-born European child. To give an example, the infant of Private Jumbe, 1st K. A. R., seen immediately after birth, had a skin colour of a pinkish white with a dark *café-au-lait* tint about it, not the dead white as seen in other young albino children; hair of head rather straight and a dark brown in colour, body and limbs covered with very light fine hair of the lanugo type; irides a liquid steel blue. Darkening of the skin occurred rapidly; it was not noticeably light after three weeks, and in three months it was an ordinary "dark" coloured baby.

Specimens of hair from the head and body having the colour as above mentioned were taken a few days after birth. The hair from the scalp was seen to contain bright yellow diffused pigment, and a fair amount of brown granular pigment. The body hairs were very fine and undeveloped, and very fine granular pigment was present.

The pigmentation, then, of the skin (irides and hair) of the native of Central Africa is a progressive process; normally pigmentation is quite marked at birth and thereafter proceeds rapidly, but in a large number does not reach, so to speak, completion till the age of possibly 16 years. This I have noticed with three or four of my younger servants who during the six years they have been with me have darkened very materially.

This production of pigment as the result of metabolism in certain cells might be likened to the production of the salivary ferments by the salivary glands. The cells are there in the skin ready for their work in the foetus just as the cells of the salivary glands are, but the latter do not begin to produce their ferments till they are needed (for the digestion of certain foodstuffs); in the same way the skin of the young foetus is unpigmented, and it is only when pigment is about to be needed that it is formed. During the few days after birth there is a call for its rapid formation and it is produced.

The stimulus for this rapid formation is probably light rays or ultra-spectral rays acting indirectly by means of a ferment, though, now that pigmentation is a racial characteristic, the tendency for the process to start and go on is represented in the germ-plasm of the parents. This inherent tendency to produce pigment is borne out by the fact that the latter was present in the hairs removed from a dermoid cyst (*vide supra*).

That depth of pigmentation is conserved as the result of such a physical cause as light I have no doubt, as the following observation will serve to show. I held a medical inspection of the 1st Battalion of the King's African Rifles the day after

part of the Battalion had returned from Somaliland. I was surprised to find when the Battalion as a whole stripped that the men who had been on foreign service were many shades darker than those who had remained in the Protectorate of Nyasaland. The Somali men were as black as any native I have ever seen, and by comparison the Nyasaland men looked like members of a yellow skinned race. The reason for this difference was that in Somaliland all fatigues had been done bared to the waist, while the men remaining in Nyasaland had not been without regimental clothing (khaki tunic or blue jersey) for years.

I obtained a similar result by keeping several turns of bandage round the chest of a native for eight weeks.

Again, I would refer to my observations on the skin colour of the penis. One may say that normally the ensheathed glans penis is not deeply pigmented; the pigmentation of the partially covered glans is proportional to the want of cover. The glans of a man who has been circumcised is invariably black. Therefore though always pigmented to some degree because skin pigmentation is a racial characteristic, the process is only completed as the result of exposure in the majority of cases. Special notice must be taken of the case of albinism of the penis; wherein, after circumcision at the age of 26 years, the skin of the glans, which was absolutely albinotic and would doubtless have remained so all the man's life had he not been subjected to operation, developed a spot of pigmentation a year later. How could this case be accounted for on the theory of structural defect*? It is a fact worthy of remark that the penis should be the commonest site to find a spotling albinism and that it should occur in close on 4 °/。 of 379 men taken at random; excluding the circumcised, it occurred in nearly 6 °/。.

In a large number of albinos we have seen there is a marked ability to form pigment (*vide* spotted albinos, and those of Class IV, who develop a certain amount of pigment in the skin but not in the hair and again those in whom pigmentation is only delayed, cases C 11, C 12 Sawali) which does not appear to point to defective structure of the skin, the explanation favoured by Pearson, but would rather suggest the action of some internal secretion. The hyperpigmentation of Addison's disease is due to the perverted action of the suprarenal glands which are commonly the seat of the destructive lesions in that disease, and it seems quite possible that normal pigmentation in negroes is due to a somewhat similar cause; light may act not directly on the skin cells but on some tissue or substance which determines the formation of pigment in the skin.

Many of the characteristics of persons living in tropical countries are akin to symptoms occurring in suprarenal insufficiency.

With regard, however, to the factor of abnormality of skin structure brought forward by Pearson, I would again refer to the conditions as seen in the sections of

* [I would venture to suggest that the very facts that the pigment developed in a spot only in this case, and that pigment occurs in localised patches only in piebalds, or in freckles and isolated locks of hair in otherwise complete albinos do indicate a differential structure; the *local* absence of a ferment—without a cause for its absence—seems to me an inadequate explanation of leucosis. Ed.]

Fig. 1.

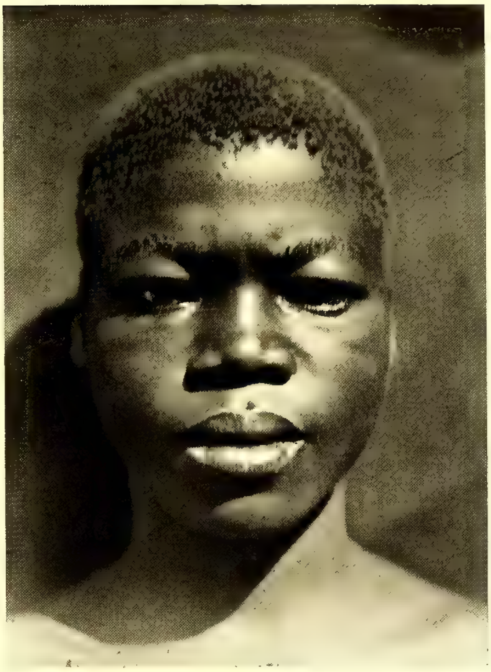


Fig. 2.

Sawali of Ntiya, Mlumbi. Case of Xanthism. See p. 346.

Fig. 3. Ouda of Checkiwere's. Case of Leucoderma. See p. 353.

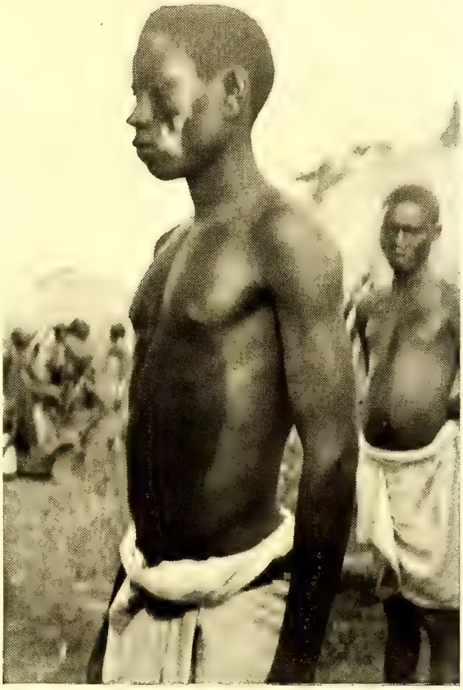


Fig. 4. Daus of Chimwai Village, Dwangwa River. Spotlings. See p. 348.

Fig. 5. Girl H., from village on Lake shore in West Nyasa. Leucotic area sharply limited along mid-line. See p. 347.

Fig. 6. Daus of Chimwai Village, Dwangwa River. Spotling. See Plate XII (4). See p. 348.

Fig. 7.

Albinotic penis not referred to in text.

Fig. 8.

Lijuni. Leucosis of glans and inner surface of prepuce after circumcision. See p. 351.

Fig. 9. Binali. Leucosis of glans and prepuce. See p. 351.

Fig. 10. Licheta, an Angoni of Dedza, Central Prison, Zomba. Case of Ziemann's "Melung." See p. 355.

Fig. 11. Matesi, an Atonga. See p. 351 (c).

Fig. 12. Jim. See p. 351 (d).

Fig. 13. Chizizu, Yao. See p. 351 (e).

Fig. 14. Lief. See p. 351 (f).

SPOTLINGS OF THE PENIS.

Fig. 15. Chiwia. See p. 351 (g).

Fig. 16. Chivia. See p. 351 (h).

Fig. 17. Naamon. See p. 351 (i).

Fig. 18. Magombo. See p. 351 (k).

SPOTLINGS OF THE PENIS.

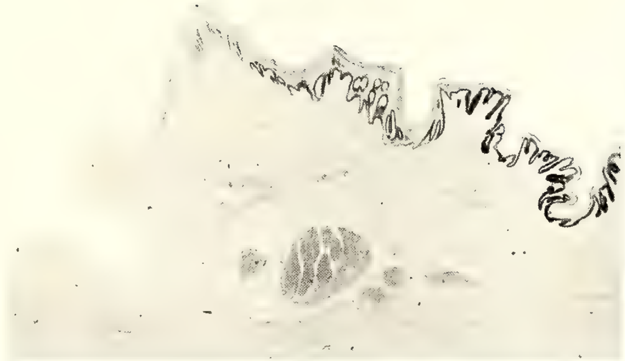


Fig. 19. ×27. Foreskin of Negro to show transition from normally pigmented skin to albinotic patch.

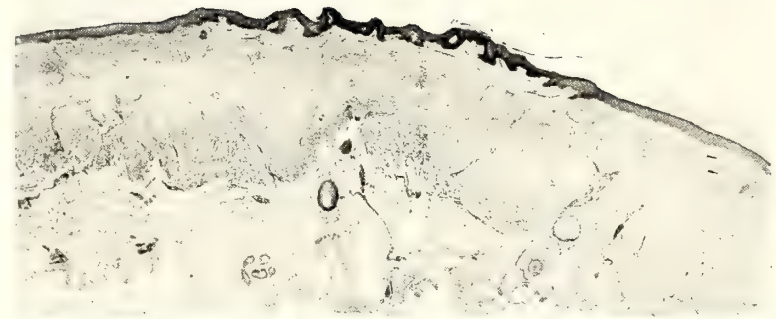


Fig. 21. ×24. Microphotograph of skin of albino showing a freckle (corrugated portion) in an albinotic area.

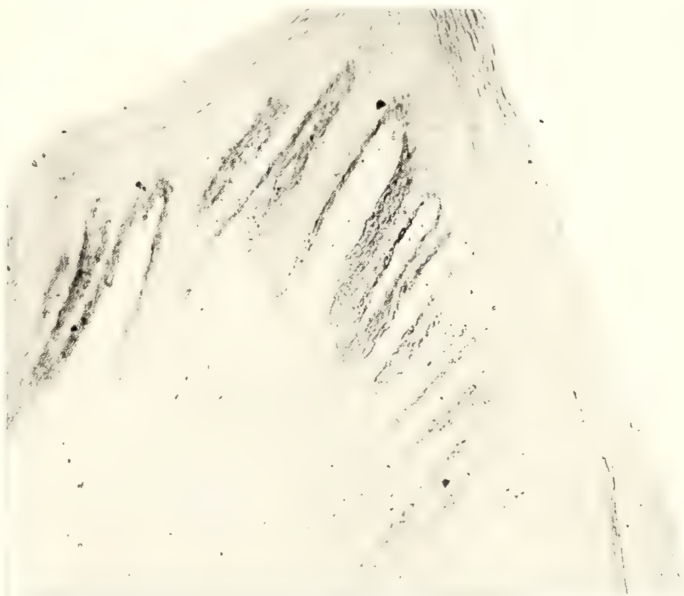


Fig. 20. ×267. The same as (19) enlarged and reversed; to show transition with gradual disappearance of melanin granules from the deepest cells of the stratum malpighii.

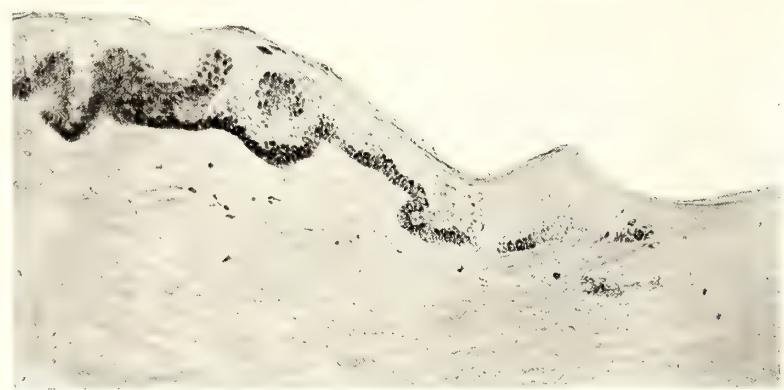


Fig. 22. ×267. Freckle in (21) enlarged to show melanin granules in deepest cells of stratum malpighii, constituting the freckle, and to indicate how these granules gradually disappear as transition is made to albinotic skin.

Fig. 23. Normal Native, Moyichandi and Sikumbiri from left to right. Grades of albinism in Nyasaland Natives. See p. 346.

Fig. 24. Chibwana, an albino with mother and sister, from Chikala. See p. 343.

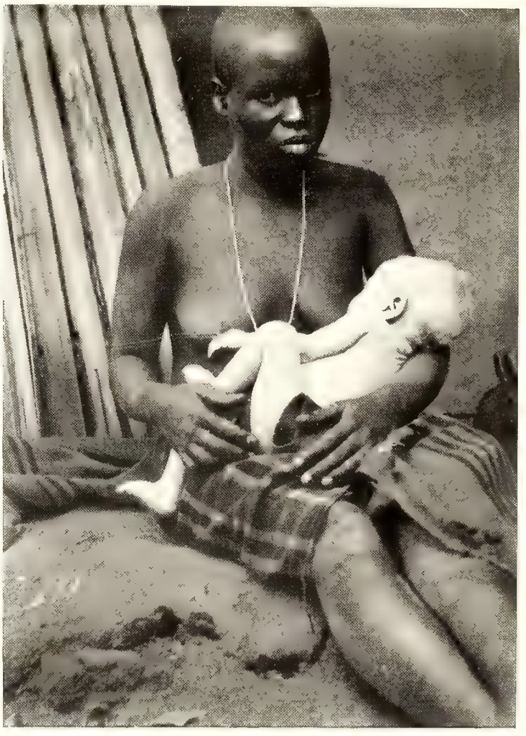


Fig. 25. Alice, an albino baby, and her mother from Chikanji's village. See p. 339.

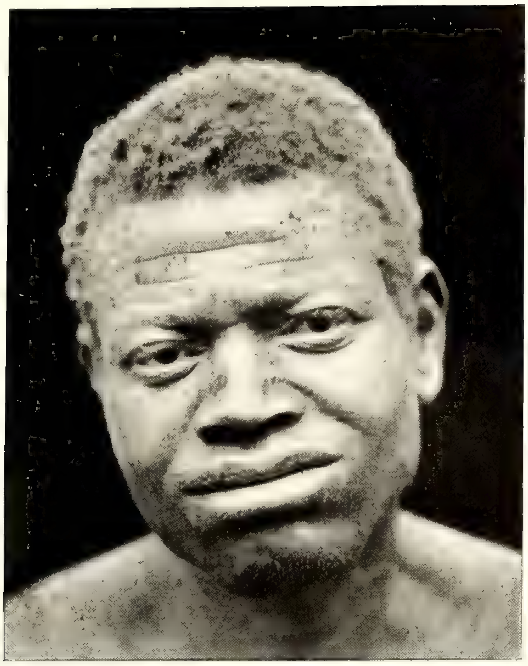


Fig. 26. Head of Pingo, an albino from Chitalu. See p. 341.

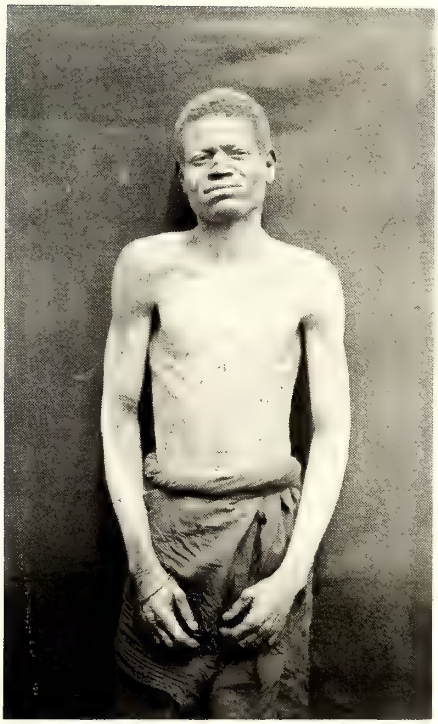


Fig. 27. Pingo, an albino from Chitalu. See p. 341.

Fig. 28. John, an albino from Mwandanas. See p. 340.

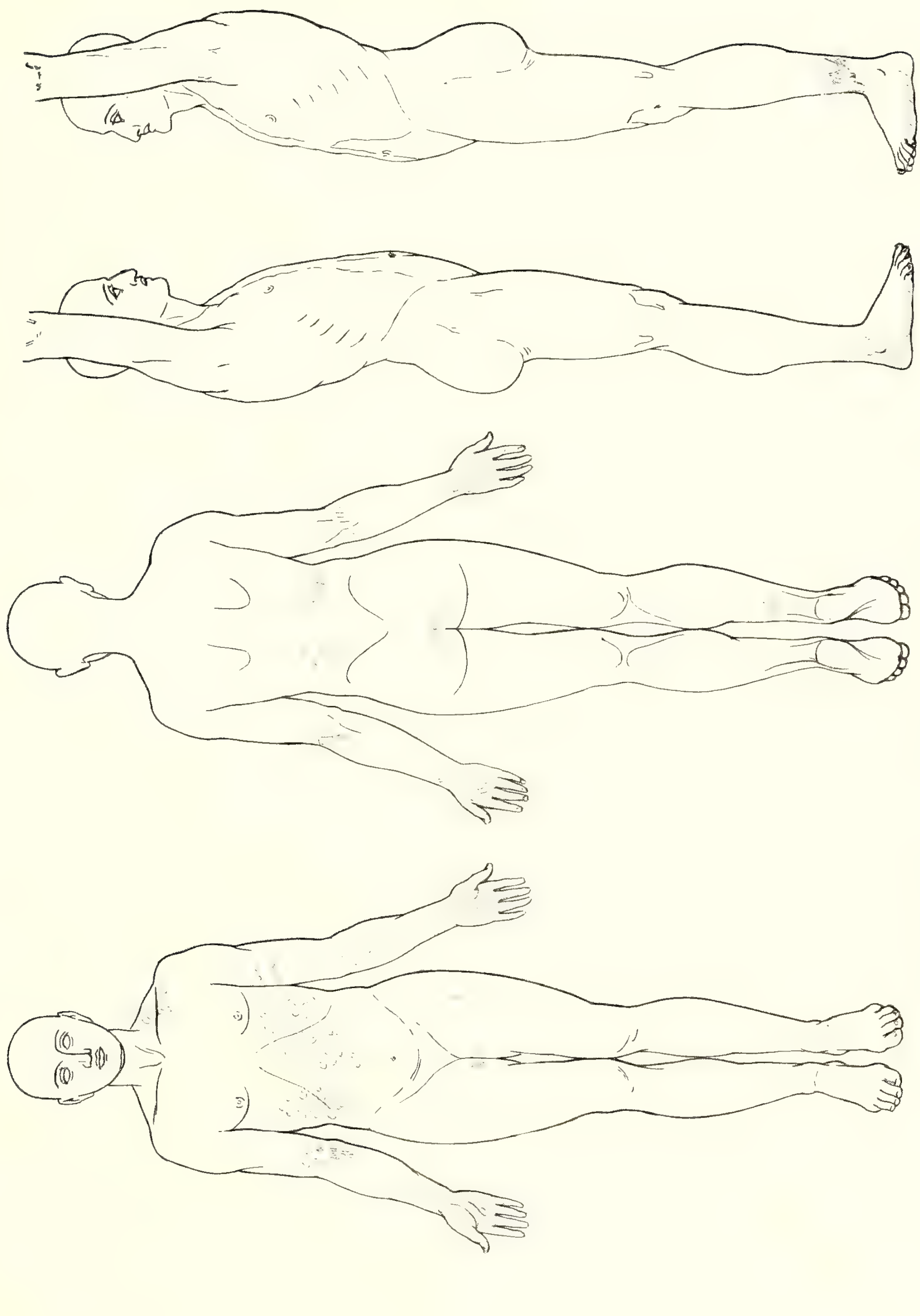
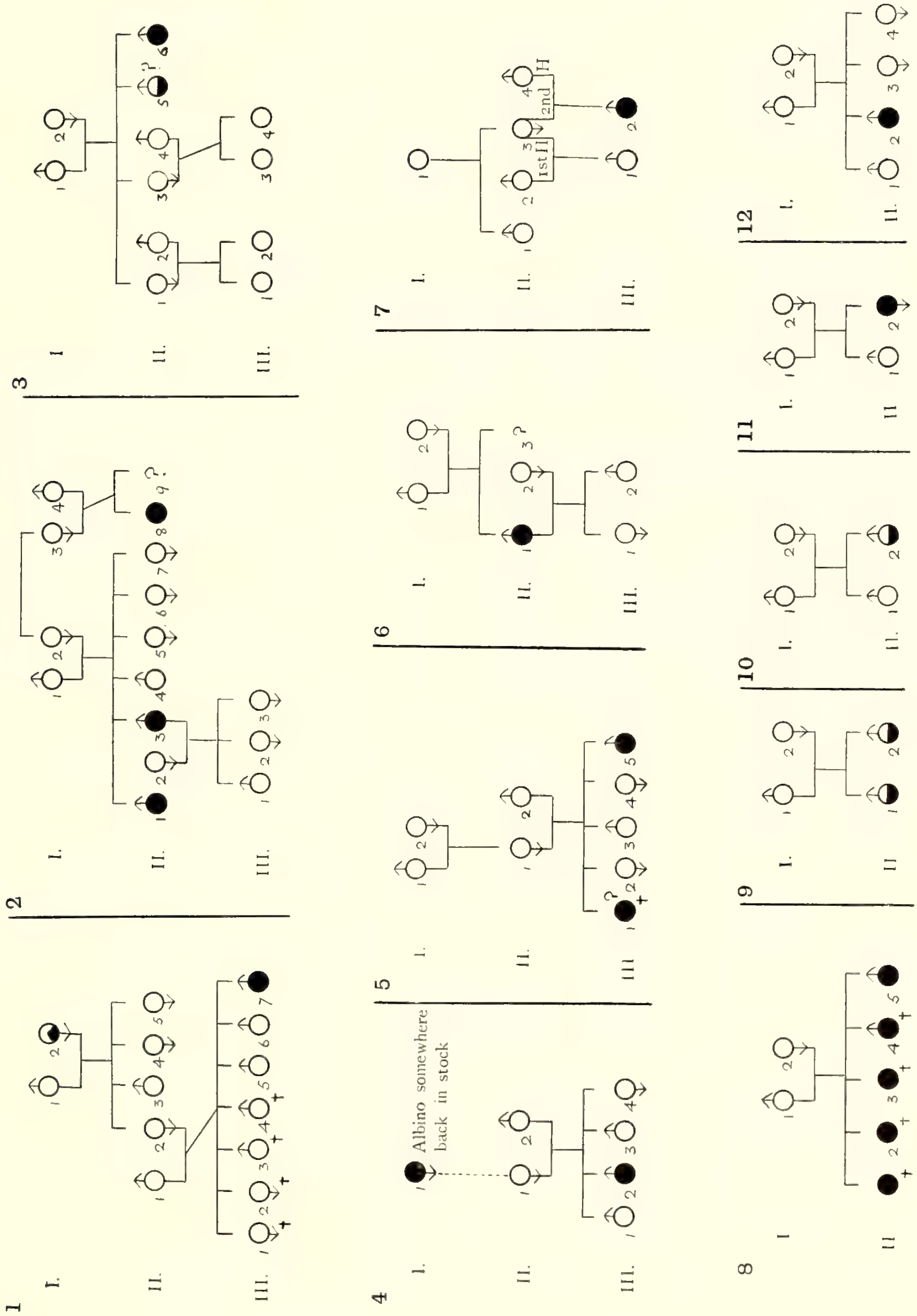


Fig. 29. Chisfufu (see p. 355). Special type of Leucoderma, Ziemann's "Melung."



Short Pedigrees of Nyasaland Albinos.

a pigmented spot from an albino, and also to the association of an unusual development of lanugo and of defective teeth with albinism.

The condition of the teeth suggests enamel deficiency, and is one which I did not meet with when examining the teeth of 1500 natives of all ages (not albinos) for statistical purposes.

One of the interesting points which has been brought to light by the investigations of Pearson and others is that in albinotic hair, while granular pigment is absent, diffuse red or yellow pigment is present, and the same is true of some kinds of red hair among Europeans. The relation of this diffuse pigment to granular pigment has not however been I think generally recognised. From the microscopical examination of the various hairs above described I think there can be little doubt that the granular pigment is formed from the diffuse pigment by some process allied to precipitation (compare hair from new-born infant child of Private Jumbe, from Mwana Ngala, from Pamandhe, also from Sawali and Moyichandi) *.

Loss of colour may be accompanied by a reversed process (compare hair from Kause, an old female lunatic).

A similar process may take place in the skin, as evidenced by the band of diffuse colouration beyond the area of granular pigment found in the section of prepuce (Lijuni).

The plan of the localised leucotic patches may be palaeogenetic as suggested by piebalds with white belly markings and flare, or there may be a segmental distribution (see case of girl H. in whom the area is limited strictly by the mid-line).

One believes therefore that pigmentation is due to an internal secretion, that light and other climatic conditions are determining factors, that a ferment may play some part, that the pigmentation of the skin is due to melanin produced in the cells of the epidermis, and that therefore structure also must be taken into account, that in albinism in its various grades and in dynamic forms of leucosis one or other or several of these factors have to be taken into account.

Addendum. One may say that in ordinary natives the colour of the iris is always some shade of brown. There is one exception. The iris in a few otherwise ordinarily dark natives may be of a greenish colour with no element of brown, a very striking feature, and known as "jeri-jeri" by the natives, who say it resembles the eye of a cat.

* [It seems difficult to establish such a conclusion until very elaborate chemical investigations of the two types of pigment, lipochrome and melanin have been made. The microscopic investigation of hundreds of samples of hair, albinotic and non-albinotic, are given in Chapter VIII of the *Monograph*, but they tell so far as much in favour of the independence as of the dependence of the two types of pigment. ED.]

ON THE EXPECTATION OF LIFE IN ANCIENT ROME, AND IN THE PROVINCES OF HISPANIA AND LUSITANIA, AND AFRICA.

By W. R. MACDONELL, Hon. LL.D. Aberdeen.

The data on which the following paper is based have been extracted from the *Corpus Inscriptionum Latinarum* of the Berlin Academy. For Rome I have examined Vol. VI. parts 1 to 4 (2), for Hispania and Lusitania Vol. II. and its *Supplementum*, also the *Ephemeris Epigraphica* VIII. and IX. 1/3, and for Africa Vol. VIII. and its three *Supplementa*.

Rome—(a) The Civil Population.

I have divided the Roman inscriptions into two classes, (*a*) those relating to the civil population, males and females being shown separately, and (*b*) those relating to the soldiers of the garrison.

Method of stating Ages. In the great majority of adults the age is given in years only, but in a considerable number the months and days are added, in a few cases even hours. Children's ages in the majority of cases are given in years, months, days and sometimes hours. Also on the soldiers' tombstones we frequently find their ages recorded to days, four of them even to hours, and amongst these we find natives of Thrace, Pannonia, etc., which points to accurate registers of births having been kept even in distant parts of the Empire. Some of the inscriptions record the ages with extreme exactness, e.g. No. 2771, 80 years 8 hours; No. 17677, plus minus 25 years 2 months 24 days; No. 20670, 86 years 1 hour.

This accuracy of statement is no doubt largely due to the requirements of astrology, which we know was widely practised in ancient Rome*; possibly also

* No. 27140, the inscription on the tombstone of a boy, who died at the age of 4 years 13 days, contains a reference to an astrologer who had evidently forecasted a long life for the child. The deceased is supposed to address the passer-by in half a dozen lines, which conclude thus:

> Non igitur, lector, lachrimes: [de]cepit utrosque
> Maxima mendacis fama mathematici.

it is in some degree due to the operation of the ordinance of the Emperor Marcus Aurelius, who was the first to introduce an official registration of births in Rome and the Provinces for the special purpose of establishing the status and age of the people when required*.

But, as I have said, in the great majority of the inscriptions the age is given in years only, and in many cases it was evidently not known accurately, probably no record of birth having been kept, or the record having been lost or destroyed. An examination of Table I will make this clear. It will be seen that after the 15th year the frequencies are heaped up on ages expressed by multiples of 5; e.g. at 25 years we find 129 deaths of males, but at 24 and 26 only 57 and 34 deaths respectively. Soldiers' ages are treated in the same way. Perhaps the explanation is that in many of these cases the age was not known exactly in years, but that the number of *lustra* of five years which the deceased had lived would be known fairly correctly, and would thus furnish the basis on which the number of years was calculated. Of course we cannot say how such a method of computation, if it existed, would be worked out in practice; if for instance it was known that no long period had elapsed since a person had completed his 6th lustrum, his age might conceivably be recorded as 30, if a considerable period had elapsed, as 35. Whatever the real explanation may be, it is evident that the practice causes a difficulty which will be obvious from an inspection of the diagrams on Figs. 1—3: the polygons showing the expectation of life in the ancient world present a jagged appearance in marked contrast to the smooth, continuously descending curves which represent the data for modern times. But these data are subjected to smoothing before publication. The difficulty might have been got over, as Mr Palin Elderton has suggested to me, by adding together the deaths recorded at ages 18, 19, 20, 21 and 22, calling these 20 (and similarly for the other groups) and then redistributing them. The same result, however, can be attained fairly accurately and with a saving of labour, by supposing the jagged polygons replaced by continuous curves, and such curves can be readily placed on the diagrams.

Date of the Inscriptions. Few of the inscriptions which record age at death give us a direct clue to the dates at which they were erected. Some that were found in the "Monumenta" of families and households of historical personages (Livia, Nero, Drusus, Marcella, etc.) can be assigned to the early Empire, and a few, mainly Christian of the 4th to 6th centuries, give the names of the contemporary consuls, and can thus be dated exactly. In one case, No. 9919, the name of the Emperor, Mauricius (A.D. 582—602), and the year of his reign are given. It is possible that scholars can date the inscriptions from the form of the letters, the spelling, etc., but as I am not aware of this having been done, I must leave the point unsettled and would suggest with some hesitation that the great mass of the inscriptions belongs to the first three or four centuries of our era.

* Marquardt, *Das Privatleben der Römer*[2], p. 86.

Status of the Deceased. A considerable number of the inscriptions record the status of the deceased, and a somewhat rough examination gives the following results for males*:

Servi	121
Vernae	164
Alumni	65
Liberti	390
Officiales and Artifices	290
Agitatores and Histriones ...	20
Patroni	64†
Ordinis equestris	19
Ordinis senatorii	8
Sacerdotes, haruspices and calatores	17
	1158

The mere fact that a person was considered worthy of a gravestone and inscription at all, would seem to indicate that the deceased did not belong at all events to the dregs of the people, and I am therefore inclined to believe that we are here dealing with a population which in status was above the very poorest and consisted largely of the slaves and freedmen, with their wives and families, of the well-to-do classes of Rome.

Expectation of Life. A reference to Table I will show how this is calculated. By way of illustration I will take Roman males. I start with a population of 4575, which I call l_0, using the actuarial symbol; of these 74 died before completing their first year, therefore 4501, or l_1, completed their first year; of the 4501, there died before completing their second year 197, therefore 4304, or l_2, completed their second year, and so on. Then the complete expectation of life at any year x is defined as

$$\frac{1}{2}+\frac{S^n{}_{x+1}(l)}{l_x}.$$

For example, at the year 1 it is

$$\frac{1}{2}+\frac{4304+4086+\ldots+1}{4501}=\frac{1}{2}+\frac{95185}{4501}=21\cdot65.$$

The values of the expectation are shown by the curves on Fig. 1, and for purposes of comparison I have added the curves showing the complete expectation of life in modern England and Wales, based on the mortality of the general population, not of healthy lives only, in the years 1891—1900. These data are of course smoothed by the Registrar-General's officers before publication.

* When the indices to Vol. VI. of the *Corpus* are published it will no doubt be possible to extend and correct these figures.

† It is worth noting that these seem to have been a long-lived class, the average age of the 64 being 62—63 years.

From Figure 1 we see that, contrary to modern experience, the curve for females always lies below that for males until about the age of 75; after that age the two curves practically coincide. For the first 20 years the male curve remains fairly level, while the female curve goes steadily down, until the difference between the two reaches a maximum about the 20th year; both remain level for the next 10 years, and then they gradually approach, the female curve rising until about 45 and thereafter falling. If our data had been ampler we should no doubt have had curves of a steady downward slope, as in the case of the African expectation.

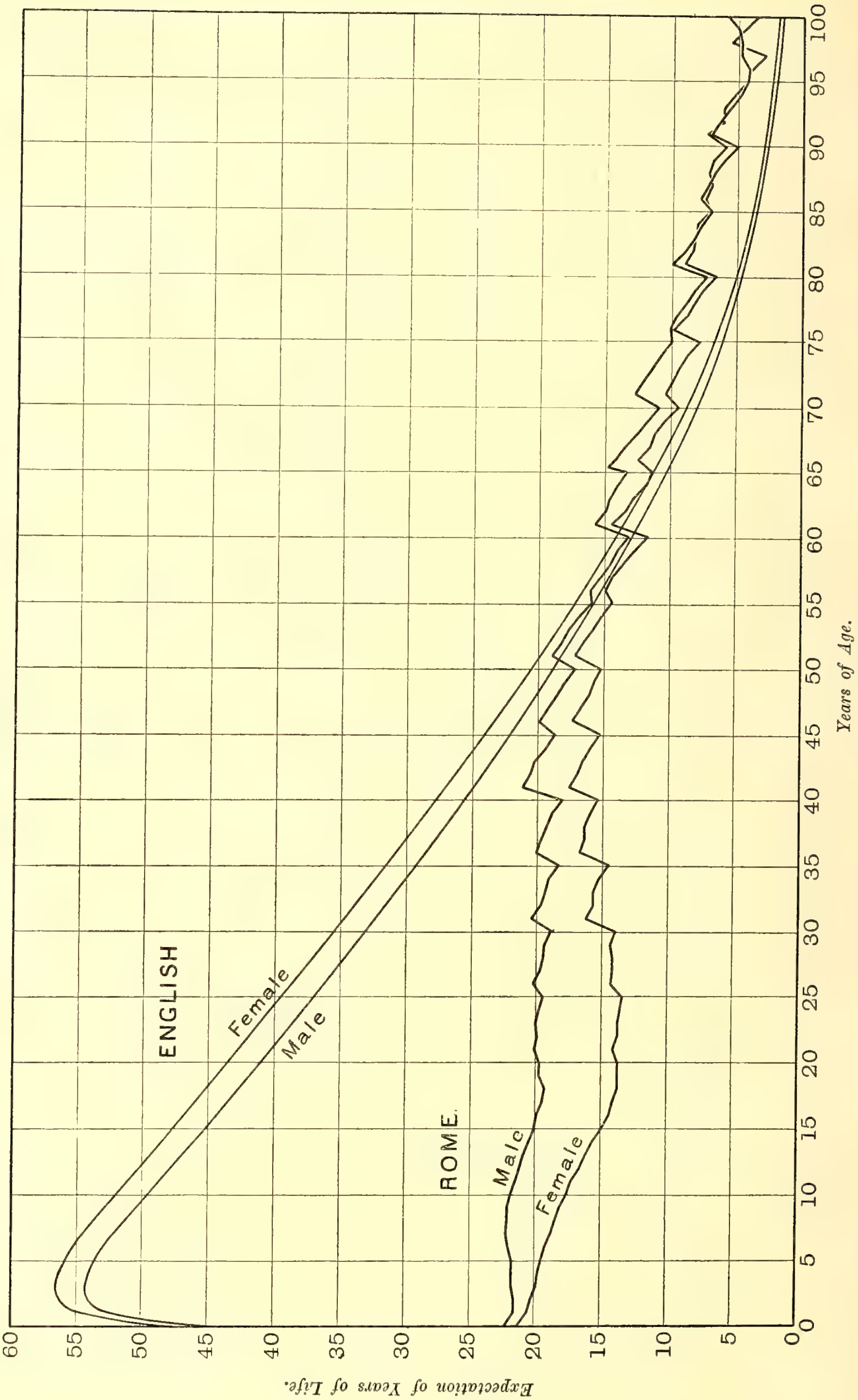
The most striking feature of the diagram is the very low expectation of life in ancient Rome, as compared with that of the present day, up to the age of 40, and especially for the first 20 years of life. This might be accounted for to some extent if we had reason to think that monuments in memory of persons who died before reaching 20 were likely to be erected in relatively greater numbers than in the case of older people. But I know of no reason to suppose that this was the case; indeed the probability seems to be the other way, for we might reasonably assume on general grounds that the older a man was the more likely he would be to make his mark in the social group to which he belonged, and be considered worthy of a tombstone and inscription. In the absence of definite information I think we may conclude that there was a relatively heavy mortality at the earlier ages, due probably to circumstances peculiar to Rome, for, as we shall see, the expectation of life was much higher in the earlier decades of age in the Provinces than in the Capital.

Between 50 and 60 ancient and modern expectations are very much the same, and after 60 the ancients had the advantage. The latter is a remarkable fact, but it is just what we should expect if the view that there was a heavy mortality amongst the young is correct; by severe selection of the young the survivors would be of specially strong physique, and would therefore show a high expectation of life.

Married Women. Out of the total of 3490 females I have noted 897 whom the inscriptions specially record as having been wives or mothers, and I find that up to the 20th year their expectation is greater than that of all women, the difference ranging from $4\frac{3}{4}$ at the age of 10 down to 1 at the age of 20; but after the 25th year it is less, the maximum difference being 2·8.

In connection with this question, it would have been interesting to know at what ages the Roman women married, but unfortunately I have found only 59 inscriptions from which this information can be ascertained. These 59 record both age at death and duration of marriage; the facts are shown in Table II. The first and last entries look doubtful, but the others seem reasonable, and would point to marriages at 10 to 20 having been quite usual. This custom of early marriages might explain to some extent the low expectation of life for women at that period, but little weight can be placed on this conclusion, as the data are so scanty.

Fig. 1. Expectation of Life in Rome.



Married Men. For the age of men at marriage the data are still scantier; I have found only 29 cases, and in these the ages at marriage run from 15 to 63, 11 of them lying between 17 and 20.

Rome—(b) Soldiers in Garrison.

The data have been collected from inscriptions 2424—3922 and 32668—33038. They record the nationality of many of the soldiers, and it appears that they came from distant parts of the Empire—Dacia, Thracia, Pannonia, Egypt, Libya and so on, as well as from various parts of Italy itself. Indeed, if we may judge from these inscriptions, the great bulk of the garrison was recruited mainly in the Danubian provinces and the north of Italy. I have therefore dealt with the soldiers separately—there are 607 whose ages are recorded—and find that as regards expectation of life they compare unfavourably with Roman males, and from 25 to 50 even with Roman married women. But no very satisfactory comparison is possible owing to the scarcity of available inscriptions.

The age of soldiers on entering the army can be found, as the inscriptions often record both age at death and length of service. Fortunately a considerable number of such inscriptions have come down to us, viz. 400 in Rome, and particulars will be found in Table III. From these figures it seems probable that the majority of the garrison soldiers joined at ages from 16 to 23, the greatest number centring round 19; but remembering the absence of precision in stating ages we can only look at the figures broadly and conclude with some probability that the hardships and risks of a military life, the change from the country to the capital, and perhaps the youthful age of the recruits had an important influence on the mortality.

Amongst the soldiers were some very old men, 19 of them ranging in age from 70 to 105; they were no doubt attached to the army in some sort of official capacity, apparently after receiving their formal discharge, if we may judge from No. 3453, which records the death of a soldier at the age of 70, and states that he was "probatus" at 22 years, served for 23, and lived 24 years 3 months and 11 days after receiving his discharge ("post missionem"). The length of service was sometimes very long; three inscriptions record 43, 51 and 55 years.

Hispania and Lusitania.

From the *Corpus* and *Ephemeris* I have extracted the ages at death of 1111 males and 885 females. I have not kept a record of the social status of this population, but from an inspection of the index in the *Supplementum* I find remarkably few set down as freedmen or slaves; it may not have been the custom to give this information, or we may be dealing with a population of higher social status than the Romans.

In very few of the inscriptions are the months and days recorded, and the same piling up of frequencies at ages represented by multiples of 5 occurs as in Rome.

There is a very marked paucity of deaths in childhood, and a relatively large number of deaths in very old age—two at 115 and one at 125 are specially noticeable. The result is that the expectation of life in these provinces, both for males and females, is far higher than in Rome for the first 30 years of life, the difference decreasing in the case of males from $17\frac{1}{2}$, and in the case of females from 13 at birth to $5\frac{1}{2}$ at the age of 30; from 30 to 45 years of age it continues to decrease, and in the period 45—75 it is about the same, the small difference being in favour of Rome, but after 75 it is again higher.

Up to 60 years of age the expectation of life of females is less than that of males, the difference varying from about 6 to 2; it is then about the same for the next 15 years, but from 75 to 95 it exceeds that of the males. From 90 onwards, however, no weight can be placed on the figures, as only five females lived beyond 90. (See Fig. 2.)

As compared with modern lives the expectation in the case of males is very much less up to 40, a little less from 40 to 50, about the same from 50 to 60, and from 60 onwards it is higher; in the case of females it is very much less up to 50, considerably less from 50 to 60, about the same from 60 to 70, and from 70 onwards it is higher. (See Fig. 2.)

Africa.

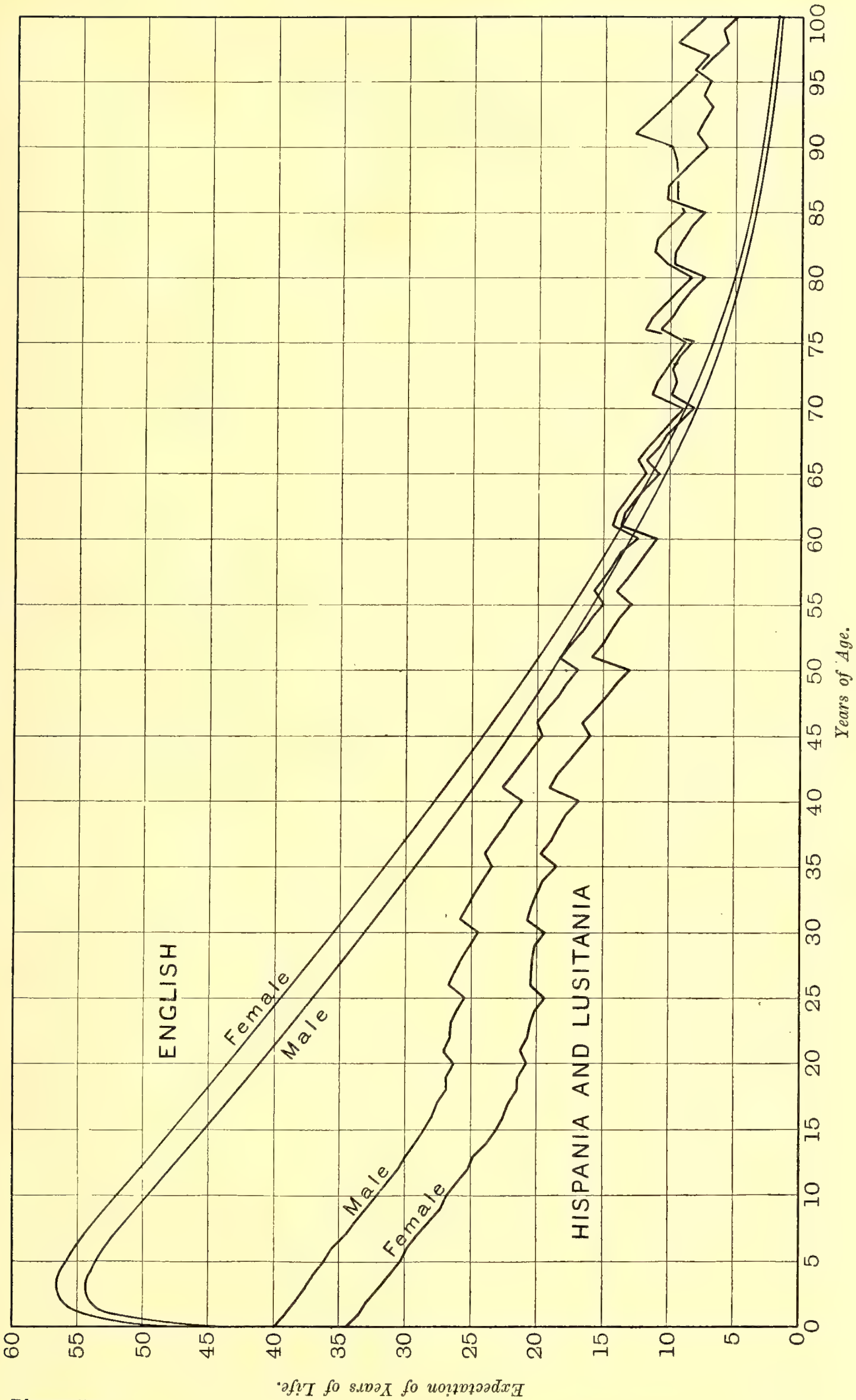
Here, fortunately, the data are more plentiful, there being 6238 males and 4459 females, and in consequence the polygons of expectation are much smoother, but before discussing the latter I have to make a few preliminary remarks which raise points of some interest.

Differences of Reading. As usual the ages are almost always given in Roman numerals (rarely in words), and when the stones have been injured or weathered or lie in inaccessible positions, naturally the numbers are difficult to read and are variously given by different observers. I have usually adopted the latest reading, or the one that seems attested by the best authority.

These discrepancies may arise in various ways; the most common in the African inscriptions are the following:

I and X in excess or defect, of frequent occurrence
I for V, e.g. XI for XV
I „ X, „ LXI „ LXX
I „ L, „ IX „ LX
V „ X, „ XXV „ XXX
X „ V, „ LXX „ LXV
V and L in excess or defect, rare.

FIG. 2. Expectation of Life in Hispania and Lusitania.



Sex. There is sometimes a difficulty, chiefly with Punic and other native names, in determining the sex, and in consequence I have omitted all doubtful cases.

Size of Families. On one inscription, No. 17463, it is stated that the deceased were the father and the mother of 12 children, from which it may perhaps be inferred that families of this size were unusual.

Record of Age. The ages are given in the vast majority of cases in years only, but occasionally months and days and even hours are added; 12794 gives the age with extreme precision: "anno uno M VIII diebus XX noctu una orabus IIII." The age of Christians seems to have been stated as a rule with greater precision than that of non-Christians: PM (i.e. plus minus) is often prefixed to the years, and months and days are relatively more frequently added; also the date of death in terms of the year of the province is often given on the Christian inscriptions.

Centenarians. It will be observed that I have recorded ages up to 132, but I have noted four inscriptions giving the ages as 140, 155, 160 and 170 years, which I have rejected as being incredible. It is a question whether the symbol for a hundred, viz. C, does not in such cases stand for *circiter*. In one case, 11902, where the C seems to have been written as a small letter above the base-line of the tens, ^{c}XXX, I have taken the age as 30. 8008, erected by a father in memory of a son, which I have not included, has VA *CV*, and the editor notes "fortasse circiter quinque." Also in 3934 the editor expands *CIR XXXV* into *cir(citer) XXXV*.

In 9106 the age of a female is given as 120 years 5 months and 25 days on the authority of Wilmanns; here the interpretation of C as *circiter* seems to be excluded by the completeness of the record and we may accept the age as correct. Perhaps I should have taken 120 as the maximum limit of age, but the few cases which I have taken above 120 affect the result very slightly.

As against the probability of those very advanced ages I may quote from 11594—the inscription on the tomb of a man who died at the age of 82 years 7 months—these words: "non digne cito vita caruisti, vivere debueras annis fere centu'," which seem to indicate that 100 was popularly considered the maximum limit of life.

Causes of Death are hardly ever assigned, but these two inscriptions are interesting: 9048, "duos una dies et *pestis* acerba abstulit hos pueros"; 18792, "O non ut meruit *pesti* vita functus est." 9050, "quae vixit sine *febribus*," would point to the prevalence of fever.

Several cases of violent death—interfectus, gladio percussus, a tauru deceptus, jugulatus, ferro petitus—have not been included.

Social Status. In the great majority of cases the social status of the deceased is not given. A considerable number were soldiers of various ranks, but their

length of service is seldom recorded; slaves are not uncommon—in two "sepulcreta familiae domus Aug. Carthag.," 331 males and 215 females have their ages recorded, and I presume most of these were slaves or freedmen, and slaves are mentioned elsewhere, but very few freedmen. There are a few pagan priests and priestesses, and a considerable number of Christians appear, their dates extending from the 3rd to the 6th century A.D.

Frequencies. There is the usual piling up of frequencies at ages denoted by multiples of 5, but the effect on expectation of life is not so marked as in the case of Rome and Hispania.

Expectation of Life. In Africa the conditions were much more favourable to longevity than in Rome and Hispania.

Up to 40 years of age the expectation for males was not more than 3 above that for females, and after 40 it was almost exactly the same. As compared with Rome, the expectation for both sexes was higher at all ages, the excess ranging from 26 for males and $24\frac{1}{2}$ for females at birth, to about 5 at the age of 60; as compared with Hispania it was higher at all ages up to 90, especially in the case of females, the excess ranging from a maximum of about 12. As compared with modern English the expectation is lower for the first 30 years of life, ranging downwards from about 8 in the case of males and about 11 in the case of females, but after 40 it is higher and goes on increasing until the end.

Conclusions. Even allowing for the scantiness of the data we may draw the following broad conclusions:

1. The evidence of the inscriptions brings out very clearly the extreme unhealthiness of ancient Rome, and on this point it agrees with the literary evidence*.

2. Leaving out of consideration ages above 75 we see that, contrary to modern experience, the expectation of life of females in Rome and Hispania was lower than that of males; in Africa there was little difference between the sexes in this respect, but the males had a slight advantage up to 40 years of age. This would point to special influences adversely affecting the health of females in the Roman empire which no longer operate in the modern world.

3. The expectation of life in Rome and Hispania, especially in the case of females, compares very unfavourably with that of modern England at the early ages, less unfavourably in middle age, but in old age it is higher, owing perhaps, as I have suggested, to severe selection in early life.

4. In Africa the expectation of life was remarkably high, especially after 40 years of age, and may be due in some degree, not only to climate and strict selection in childhood, but also to the considerable influx of colonists from Italy which

* Friedländer, *Sittengeschichte Roms*[8], I. 39 ff.

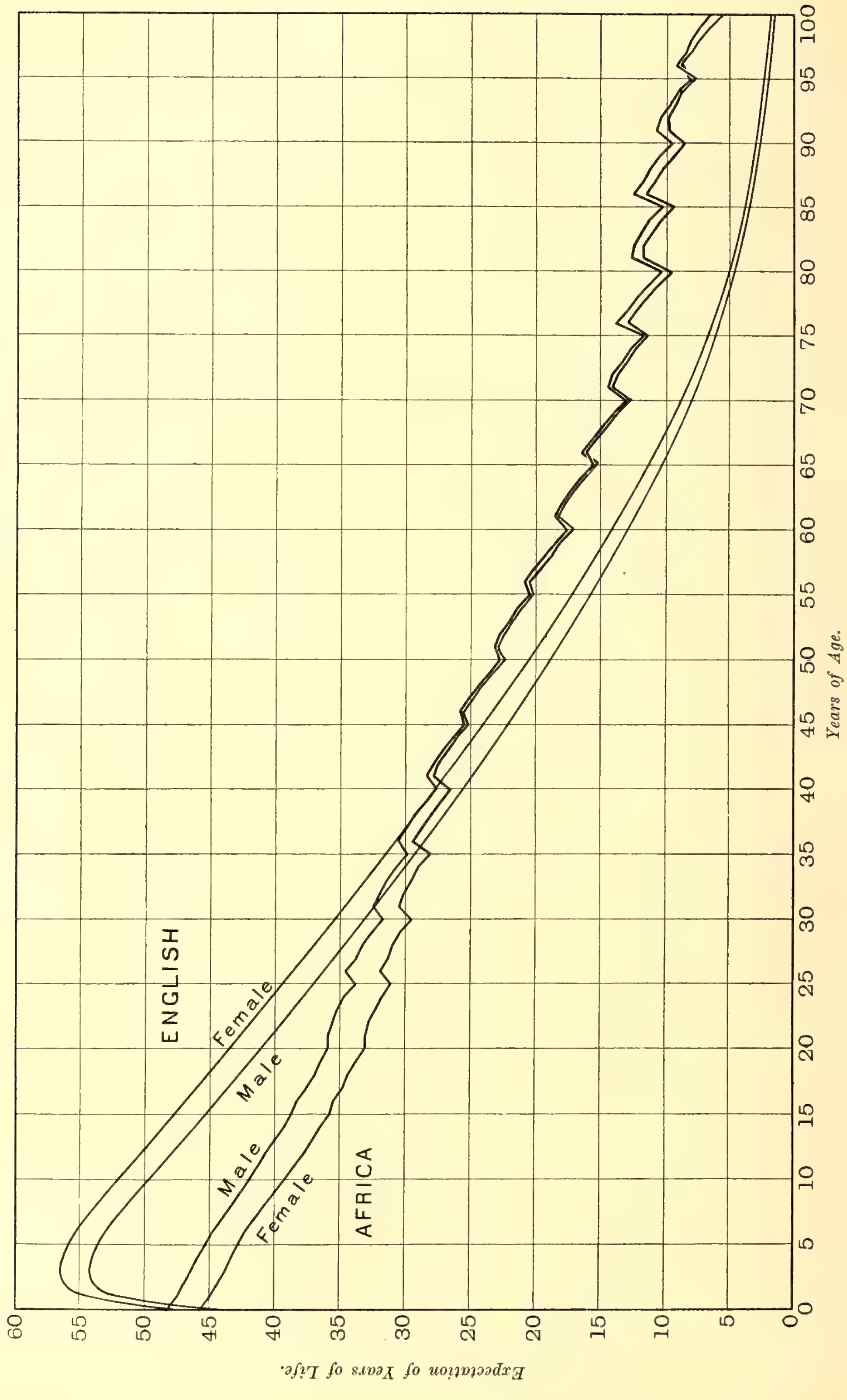


Fig. 3. Expectation of Life in Africa.

took place in the time of Caesar and the early Empire*—for we may reasonably suppose that the colonists were a selected class, possessed of vigour and enterprise, qualities which would be inherited more or less by their descendants—and also to the fact that a large proportion of the population must have been engaged in the healthy occupation of agriculture, for we know that Africa rivalled Egypt as a source of supply of grain for the markets of ancient Rome†.

Note. After preparing the material for the Roman and Spanish part of this inquiry I found from the *Berliner Philologische Wochenschrift*, Vol. 24, p. 604 (7 May 1904) that Professor A. de Marchi published a paper in the *Atti dell' Accademia di Padova*, 1903‡, on the average duration of human life in antiquity based on the inscriptions in Vol. VI. of the *Corpus*. He seems to have gone over the same ground as I have done as regards Rome, and has also dealt with *Gallia Cisalpina*, where he has found a much smaller mortality amongst children than in Rome; in this respect it is interesting to find another province agreeing with Hispania and Lusitania and Africa. I regret that I have not seen Professor de Marchi's paper; coming from so distinguished a scholar it carries an authority to which the present essay cannot lay claim.

* Even before Caesar's time, ever since Rome subdued Carthage, a sporadic emigration of Italians into North Africa had been going on (J. S. Reid, *The Municipalities of the Roman Empire*, Cambridge, 1913). Professor Reid points out, however, that "emigration on anything like the modern scale, or even on the ancient Greek scale, was unknown to the Romans," and with regard to Roman Africa he writes: "The impression that must be produced by a scrutiny of the evidence is that but a minute fraction of the population could claim Roman descent." (See page 316 of his book.)

† Mommsen, *Römische Geschichte*[3], v. 647, 651.

‡ The reference appears to be incorrect as no such paper could be found in the *Atti* in a careful search of the volumes at the British Museum.

TABLE I. *Number of Deaths at each Year of Life in Rome and the Provinces.*

Years completed at Death	Rome		Hispania and Lusitania		Africa	
	♂	♀	♂	♀	♂	♀
0	74	46	4	2	32	16
1	197	127	6	6	60	44
2	218	95	10	4	56	41
3	185	131	10	5	63	46
4	188	117	5	6	49	39
5	202	99	9	13	56	52
6	176	91	3	6	39	27
7	142	99	5	5	42	30
8	132	89	6	3	46	21
9	107	78	8	13	36	28
10	91	81	8	10	53	24
11	84	64	4	7	52	29
12	93	68	14	16	42	41
13	79	63	9	4	53	44
14	72	54	12	5	47	34
15	92	75	14	18	86	71
16	84	91	21	19	46	41
17	94	91	17	15	51	56
18	152	153	31	31	81	50
19	106	109	16	10	60	34
20	138	158	58	49	142	110
21	76	62	15	13	95	77
22	96	117	25	23	78	48
23	68	75	18	18	69	56
24	57	70	15	7	23	28
25	129	152	62	60	238	192
26	34	64	13	21	32	41
27	55	64	14	22	69	64
28	64	73	10	18	39	38
29	31	34	9	7	21	16
30	153	157	60	51	229	196
31	16	18	10	10	65	56
32	38	38	10	12	52	34
33	35	22	9	8	41	48
34	19	12	7	3	16	8
35	131	100	35	39	222	208
36	23	18	8	3	34	24
37	28	25	8	8	46	27
38	24	16	6	7	24	28
39	10	10	5	1	18	16
40	146	68	55	50	229	202
41	14	7	5	6	46	42
42	20	11	4	2	39	20
43	4	7	7	3	30	28
44	9	4	4	2	17	9
45	64	51	30	24	167	122
46	9	5	4	—	32	14
47	9	5	4	3	40	18
48	10	10	9	2	20	15
49	7	4	2	—	15	7
50	69	37	49	51	187	131
51	7	4	1	2	72	34
52	10	3	4	4	33	23
53	4	3	6	4	48	25
54	2	3	3	1	15	9
55	27	16	31	22	168	102
56	5	5	6	2	27	10
57	7	1	5	3	31	17
58	12	1	4	4	21	18
59	3	1	—	2	9	5
60	75	34	53	31	245	194
61	4	—	9	4	60	35
62	11	3	4	2	28	14
63	7	—	5	1	29	23
64	6	4	3	—	6	2
65	30	16	25	13	188	137
66	3	1	4	1	27	14

TABLE I.—(*continued*).

Years completed at Death	Rome		Hispania and Lusitania		Africa	
	♂	♀	♂	♀	♂	♀
67	3	4	1	2	23	15
68	4	1	2	—	20	7
69	1	—	—	—	7	5
70	43	14	48	17	277	179
71	4	3	6	3	67	42
72	5	3	3	6	27	14
73	4	2	3	2	42	15
74	2	—	3	—	8	4
75	16	16	25	14	249	182
76	6	—	1	1	22	7
77	4	1	3	—	18	11
78	4	2	1	—	14	3
79	1	—	—	—	2	6
80	39	12	23	6	264	155
81	—	1	4	3	62	30
82	3	2	2	1	22	13
83	3	2	2	—	16	12
84	3	—	—	—	6	1
85	13	3	14	2	173	102
86	3	1	2	1	12	7
87	3	2	—	1	16	9
88	2	1	—	1	9	7
89	—	—	—	1	9	1
90	17	3	5	2	95	58
91	1	—	1	—	33	12
92	1	1	1	—	14	5
93	1	—	3	—	18	5
94	2	—	1	—	3	—
95	4	1	3	—	70	43
96	5	—	—	—	6	4
97	2	3	3	—	14	9
98	2	—	—	1	5	1
99	2	—	—	—	6	2
100	1	1	2	2	46	39
101	—	—	2	—	19	33
102	—	—	—	—	6	4
103	1	—	—	1	7	10
104	—	—	—	—	1	—
105	2	—	—	—	39	20
106	—	1	—	—	3	1
107	—	—	—	—	5	1
108	—	—	—	—	—	1
109	—	—	—	—	1	—
110	—	—	—	—	21	8
111	—	—	—	—	3	2
112	—	—	—	—	1	—
113	1	—	—	—	1	—
114	—	—	—	—	1	—
115	—	—	1	1	7	9
116	—	—	—	—	—	—
117	—	—	—	—	—	1
118	—	—	—	—	—	—
119	—	—	—	—	—	—
120	—	—	—	—	9	3
121	—	—	—	—	1	1
122	—	—	—	—	—	—
123	—	—	—	—	—	1
124	—	—	—	—	—	—
125	—	—	1	—	4	4
126	—	—	—	—	—	—
127	—	—	—	—	1	—
128	—	—	—	—	—	—
129	—	—	—	—	—	—
130	—	—	—	—	—	1
131	—	—	—	—	1	—
132	—	—	—	—	—	1
Totals	4575	3490	1111	885	6238	4459

TABLE II.

Age of Roman Women at date of Marriage.

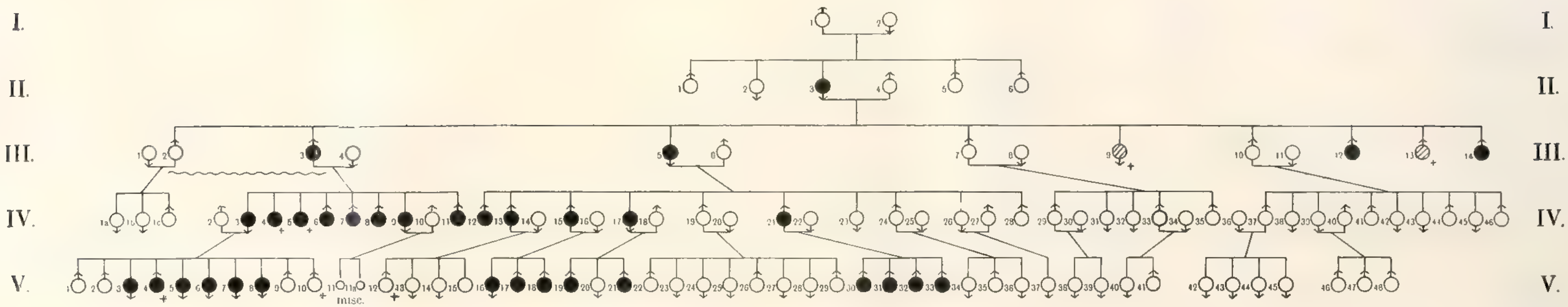
Age	Number of Women	Age	Number of Women	Age	Number of Women
7	1	*17*	1	*28*	1
10	3	*18*	4	*30*	1
12	7	*19*	2	*31*	1
13	6	*20*	6	*32*	1
14	6	*21*	2	*56*	1
15	8	*24*	1		
16	5	*25*	2		59

TABLE III.

Age of the Soldiers of the Roman Garrison on Entering the Army.

Age	Number of Soldiers	Age	Number of Soldiers	Age	Number of Soldiers
15	2	*24*	8	*33*	1
16	15	*25*	12	*38*	1
17	28	*26*	7	*43*	1
18	72	*27*	9	*45*	1
19	59	*28*	1	*47*	2
20	91	*29*	2	*52*	1
21	31	*30*	3		
22	29	*31*	2		400
23	18	*32*	4		

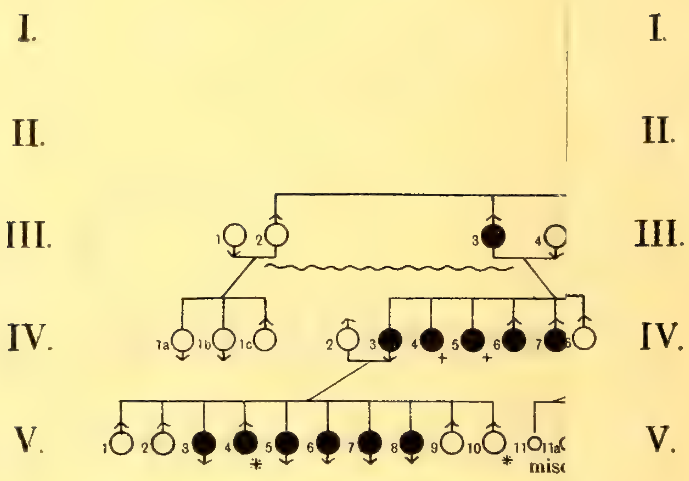
PEDIGREE OF THE *S.* FAMILY WITH HEREDITARY "LOBSTER CLAW."



$*$ = Still-born.

+ = Died young

After this pedigree was engraved the Rector of N. wrote to the Editor stating that he had found at last the entry of *W. S.* (III. 2) in the baptismal register, two years *later* than that of III. 3, so that doubt as to order of birth of III. 2 and III. 3, expressed by the wavy line in the diagram, is now removed.



* = Still-born.

After this pedigree was engraved the Rector of h

ON THE INHERITANCE OF THE DEFORMITY KNOWN AS SPLIT-FOOT OR LOBSTER-CLAW.

(SECOND PAPER.)

By GEORGE McMULLAN, M.D. and KARL PEARSON, F.R.S.

In 1908 Pearson discussed the pedigree of a certain Lobster-Claw or Split-Foot Family in this Journal*. Pearson's information was obtained from several members of the family, but his photographs and skiagrams were of one particular section of it. McMullan, unacquainted with Pearson's paper, read a paper on the same family before the Reading Pathological Society in 1910. He worked principally with a different section of the family and his photographs and skiagrams did not cover the same individuals as Pearson's. Hearing of the latter's paper, McMullan placed his material at the disposal of the Editor of this Journal and the present supplementary memoir is the result. During the five years that have intervened since the first paper several additions have been made by birth to the family; thus while Pearson dealt with 25 deformed individuals the present account contains 33†. Further, somewhat fuller particulars have been found of the two earliest generations. The order of birth of Generation III. and further the total number of brothers and sisters were given differently by different members of the family to Pearson and to McMullan. These discrepancies led Pearson to make some investigation into the church registers, but they disclosed a rather chaotic state of affairs. Some facts, however, resulted. *Ann J.* (II. 3) was married to *Francis S.* (II. 4), at N. in the Chilterns in 1837, both being unable to sign their names. It was supposed that III. 2, 3 and 5 were born before the *S.'s* migrated to W. The names of most of the other children in Generation III. are to be found at about the proper dates in the baptismal register of W. But not until the birth of *John S.* (III. 13), who must have died under two years of

* *Biometrika*, Vol. VI. p. 69, 1908.

† Besides six additional births of deformed children since the date of Pearson's record, IV. 9 has recently stated that her parents had two deformed children who died as infants, IV. 4 and IV. 5. The statement is now confirmed by IV. 3.

age,—for his brother born in 1866 is also *John S.*—do the names *Francis* and *Ann S.* reappear. The births of *William S.* (III. 2), *James S.* (III. 3), and *Elizabeth S.* (III. 5) are fixed by the registers of N., but another *James S.*, and an *Eliza Ann S.*, *Charles S.* (III. 7), *Emma S.* (III. 9) and *Thomas S.* (III. 10) are attributed in the registers of W. in four cases to *Thomas* and *Frances S.* and in one case to *William* and *Ann S.*! There was thus very great carelessness in the preparation of the baptismal register, or our illiterate couple frequently changed their names, or there were other families of the same surname and with children with the same Christian names in the parish, although there is no other trace now of their former existence. Ultimately, therefore, we are again thrown back on the statements made by the two survivors of Generation III., who, it may be noted, did not give their parents names correctly, or on those of the grand-daughter IV. 3, who has, after an interval of five years, given almost the same account of Generation III. as she did in 1908, and on both occasions referred to III. 2, III. 9 and III. 13, who seem substantiated by the registers, although they were not referred to by III. 7 and III. 14. Probably Generation III. is now in fairly correct order, although it is possible that two further children—who must have died young—namely, another *John* (1848) and another *Thomas* (1852), really belonged to this pair, who were married in 1837 and were still having children in 1866. Of course it is not possible to ascertain whether such children or III. 9 and 13 really were deformed or not, although IV. 3 asserted that III. 9 was normal, presumably on the authority of her father, III. 3.

In the following account a pedigree number with P. P. affixed refers to the index number of the corresponding individual in Pearson's Pedigree in the earlier paper. The accounts given there are not reproduced unless there are additions to be made to them.

Account of Pedigree.

I. 1 and I. 2 are now given and given as normal, because of the following legend, recently unearthed, which attributes the origin of the deformity to "maternal impression."

I. 2 the great-great-grandmother of the latest generation being pregnant committed a theft; when charged with the offence she denied it and further invoked her Creator to let her child when born have no fingers to steal with, should she be guilty. In due course the infant was born without fingers or toes and the woman confessed to the theft lest a greater evil should befall her. All which account is very picturesque but somewhat unconvincing. As far as one can gather, the woman was well advanced in pregnancy; and no explanation is suggested of the process by which the digits were separated, nor of their ultimate fate. Perhaps the appeal is to the miraculous rather than to a maternal impression. But the legend is so far of interest that it seems unlikely that the mother or father of II. 3 were deformed, otherwise, there would have been no necessity to account for the appearance of the deformity in II. 3. Pearson gave

II. 3 (P. P. I. 2) one brother and one sister (P. P. I. 3 and 4), but was unable to find any real evidence as to their condition. McMullan's informants provide no less than three brothers and one sister, II. 1, 5 and 6 and II. 2 respectively, all stated to be normal. But no evidence is forthcoming as to their names, place of origin, residence or descendants. *Ann J.* (II. 3) is reported to have been a servant in a farmhouse at N., and the legend suggests that her mother was possibly of the same class as some of her more recent descendants, which have included tramps, wandering labourers and paupers. It seems almost impossible therefore to trace any of this *J.*-family now; some years ago Pearson made, without result, inquiries of residents and medical men in the neighbourhood of N., both as to the family and as to the existence of any similar deformity in the district. No stress ought to be laid on the numbers in Generation II., they are probably not exact. Of *Francis S.*, II. 4 (P. P. I. 1), we know nothing beyond his name and the fact that he could not write it. There is no evidence whatever to suppose him in any way of the same stock as II. 3*. It is clear that II. 3 could not, in Mendelian language, have been either a pure recessive (*RR*) or a pure dominant (*DD*). In the former case she would have had no affected children, in the latter case all affected children. She must be looked upon as an impure dominant (*DR*). Thus every normal in the pedigree must be looked upon as an (*RR*), and consequently since every mate is a normal, all the affected can only be (*DR*)'s.

In Generation III., we have at least nine children born to *Ann J.* (*Mrs S.*), perhaps there may have been one or two more who died young and without possibility of any record of deformity surviving. Of these nine, four were affected, three are known to have been normal, one is reported normal (III. 9), and the nature of the other (III. 13) is certainly unknown. This gives a quite reasonable Mendelian 50 %.

In Generations IV. and V., where we know the individuals much more accurately, there are eight matings of affected (*DR*) with normal (*RR*). Of these that of IV. 9 and 10 contributes nothing as there have been only two early miscarriages. The remaining seven matings have so far produced 41 offspring, of whom 28 have been affected, nearly two-thirds. The ratio is thus nearly 2 : 1 and not 1 : 1. The odds against such an excess as 28 to 13 are about 100 to 1, and they tell very heavily against any simple Mendelian theory applying to this deformity.

Thus far the normal members of the stock *when mated with outside normals* have bred only normals. This is no convincing argument for Mendelian theory, because in no case has an intra-stirp normal been mated either with another intra-stirp normal or with an abnormal. Both such matings would have to be considered before we could assume that an intra-stirp normal is as free from taint as an extra-stirp normal. Further the fact that three out of the seven matings have given nothing but deformed offspring is most remarkable and almost impossible on any purely Mendelian theory. If III. 3, IV. 15 and IV. 21 had been pure dominants,

* IV. 3 distinctly states that there was no consanguinity.

(*DD*)'s instead of (*DR*)'s, this result would have occurred. While if III. 5, IV. 3, IV. 13, IV. 17 had been (*DR*)'s their 26 offspring would have given 13 affected and 13 unaffected as they actually do. It appears fairly clear that III. 3, IV. 15 and IV. 21 were in some manner far more prepotently affected with the taint than the other four. But even if we examine the other four, IV. 3 and IV. 13 seem to be affected in quite different degrees; thus not only on the average, but individually, the seven families are very improbable on the Mendelian hypothesis; they suggest a graded intensity of taint in the germ-plasm of the different affected members of the stock. But the moment such relative prepotency of the individual is admitted, Mendelian segregation in definite ratios falls to the ground. A still more difficult point is the definition of what a Mendelian "unit" character is in such a case as the present. McMullan's fuller descriptions completely confirm the very great variation in the deformity, which as Pearson pointed out in 1908 could affect any number from 10 to 60 bones. But it is not only the bones that are affected. The usual type is a claw-like foot or hand, the digit or digits remaining being bent round at an angle to the remainder of the hand or foot. This is very manifest in the bulk of the skiagrams and photographs of the present and the earlier paper. But the case of V. 8, born since the earlier paper, and several photographs in this paper show that a single *straight* digit is a possibility.

Ann J. II. 3 (P. P. I. 2). *Mrs Francis S.* Married to *Francis S.* at N., April 15, 1837. There is nothing to add to the description of the deformity given in this case by Pearson.

James S. III. 3 (P. P. II. 6). Born March 4, 1838. Both hands and feet deformed. Pearson's and McMullan's informants agree as to feet both presenting 1st and 5th toes only. But Pearson's informant (IV. 3) says there were two fingers on right hand, one finger and thumb on left; while McMullan's informant says there was only the fourth finger on both hands. The difference between the right and left hands stated positively by the eldest daughter must, we think, stand.

William S. III. 2 (P. P. II. 2). Born February 10, 1840. Normal. Year of death not found. Three normal children IV. 1 *a—c* (P. P. III. 1—3).

Jane R. III. 4, a native of C., was quite unrelated to her husband, III. 3.

Elizabeth S., afterwards *Mrs Wh.* III. 5 (P. P. II. 3), dead. Born April 23, 1842. Both McMullan's and Pearson's informants agree that she had normal hands; the former reports both feet as having only their 5th toes, the latter as having the 1st and 5th toes.

—— *Wh.* III. 6 (P. P. II. 4). Husband of III. 5; he was unrelated to his wife, but is now reported to have had a deformed foot, quite different from those of his wife, and probably to judge by the description a club-foot.

George S. III. 12 (P. P. II. 10) (dead). Date and locus of birth? no information beyond that already given by Pearson that both hands and feet were deformed.

John S. III. 14 (P. P. II. 3), seen by Pearson's field worker and again by McMullan*. Born, 1866. Scavenger, unmarried; both hands and feet affected.

Feet. Examination refused, but said he had the 5th toe only present on each.

* Bones present and those absent in this as in other cases made out by palpation. Such a method leaves something to be desired, but many of the affected persons feel their affliction keenly and are not willing to have it further investigated even if X-ray apparatus were available, which it is not in country districts.

Hands. Carpus apparently normal on both right and left; metacarpals: 1st, 3rd, 4th and 5th present on right hand, 2nd absent or rudimentary; 3rd, 4th and 5th present on left hand; 4th finger three phalanges present on both hands. Thumb on right only, rudimentary, one phalanx.

—— *S.* *Mrs R.*, IV. 3 (P. P. III. 21). McMullan's informant gave no additional information, and the skiagrams in the first paper fully describe the case.

IV. 4 and 5. Two children, sex unstated, reported recently by IV. 9, said to have died as infants and to have been deformed.

George S. IV. 6 (P. P. III. 22). Dead, unmarried. Hands and feet deformed; deformities irregular, described by Pearson's informant; they involve polydactyly: see *loc. cit.*

Richard S. IV. 7 (P. P. III. 23). Inmate of workhouse, unmarried. Hands and feet deformed.

Hands	*Right*	*Left*
Carpus.	Trapezium probably absent.	Normal.
Metacarpals.	2nd, 3rd, 4th and 5th present.	1st, 2nd, 3rd, 4th and 5th present, 1st rudimentary.
Phalanges.	Mid-finger, one rudimentary phalanx. 4th finger, three phalanges.	First finger, one rudimentary phalanx. Ring finger, one fairly well-developed phalanx. Little finger, three phalanges.

Feet. Both show the same deformity.

Tarsus. Probably middle and external cuneiform bones absent.

Metatarsals. 1st and 5th only present.

Phalanges. 1st toe, two fused phalanges. 5th toe, three fused phalanges (i.e. no moveable joints are demonstrable). A web of skin is present between the two digits.

For photographs of hands and feet: see Plate XXVII, Figs. (xv) and (xvi).

Francis S. IV. 8 (P. P. III. 26). Farm labourer. *Hands* said to have one finger only, the 4th, present.

Feet said to resemble last case.

Annie Eliza S. (*Mrs F.*). IV. 9 (P. P. III. 24). Married, no living children; has had two abortions at the 4th month (V. 11—12). Has both hands and feet deformed.

Hands	*Right*	*Left*
Carpus.	Normal.	Normal.
Metacarpals.	Normal.	Normal.
Digits.	Thumb, one very rudimentary bone.	Thumb, short otherwise normal.
	Index finger absent.	1st finger, one small rudimentary phalanx.
	Mid and ring fingers united, one nail only present, and bones partially fused to each finger, three phalanges.	Mid finger, very long, three phalanges. Ring finger dichotomous at 2nd joint.
	4th finger normal.	4th finger normal.

Feet. Both show a similar deformity. Presence of middle and external cuneiform bones doubtful. 1st and 5th metatarsals only present. Phalanges, normal in number, but ankylosed.

For photograph of hands: see Plate XXV, Fig. (vi); and for that of feet: see Plate XXVII, Fig. (xviii).

Herbert S. IV. 11 (P. P. III. 27). Farm labourer, unmarried. Hands and feet deformed.

Hands	*Right*	*Left*
Carpus.	Apparently normal.	Apparently normal.
Metacarpals.	1st, 3rd, 4th and 5th present.	1st, 2nd (probably present but small and rudimentary), 3rd, 4th and 5th fairly well developed.
Digits.	Thumb, rudimentary, one small phalanx. Ring and 5th fingers united by thick fleshy web. Three phalanges to each.	Thumb rudimentary, one small bone. Middle and ring fingers fused, one nail. Three phalanges to each. 5th finger normal.

Feet. Show a deformity similar to *Eliza S.* (IV. 9).

For photographs of hands and feet: see Plate XXVI, Figs. (ix) and (x).

Wm. Wh. IV. 12 (P. P. III. 4). Dead, unmarried. Hands and feet said to have been deformed. Said to have had only the little finger and 5th toe present.

Alfred Wh. IV. 13 (P. P. III. 5). Married. (Four normal children, V. 12—15.) Hands and feet said to be deformed. Hands said to be broad with 5th digit present on each and a thumb on right hand. Feet to have 5th toe only present.

Chas. Wh. IV. 15 (P. P. III. 7). Married, whereabouts unknown (has three deformed children). Hands and feet in this case said to be affected; in what manner not specified.

Eliza Wh. (*Mrs Sm.*). IV. 17 (P. P. III. 9). Recently an inmate of Claybury Asylum suffering from confusional insanity. Discharged recovered. Hands and feet deformed: has had three children, two male deformed and one female normal. One of the male children and the female child are dead. Stated when in asylum that she had had a fourth child with deformed hands and feet when sixteen years old, but statement doubtful.

Hands	*Right*	*Left*
Carpus.	Normal.	All the bones appear to be present though the trapezium and trapezoid are not well developed.
Metacarpals.	1st very rudimentary. 2nd badly developed. 3rd, 4th and 5th normal.	3rd present, rudimentary. 4th and 5th well developed but showing a measure of deformity.
Digits.	One, the 5th, alone present, shows ankylosis.	A small bone lies between the heads of the 4th and 5th metacarpals, possibly a rudimentary digit. The 5th digit present and badly developed.

Feet. Tarsus. In the case of the feet the deformity is difficult to determine, both however show either absence or very imperfect development of the middle and external cuneiform bones.

Metatarsals. The 1st and 5th alone are present in the left foot while the right shows these two bones, and also a rudimentary 4th metatarsal.

Digits. 1st and 5th alone present, the number of bones in each is normal, but much deformity is shown, i.e. ankylosis, dislocations, etc.

Skiagraphs, hands and feet, Plates XXIII—XXIV; photograph, Plate XXV, Fig. (v).

Our best thanks are due to Dr Robert Jones of Claybury Asylum for acquainting us with the facts regarding this case, and providing a photograph and skiagraphs of the deformities.

G. Wh. IV. 19 (P. P. III. 11). Normal, married, with eight normal children, V. 22—29 (P. P. IV. 8—14, and one additional birth).

Thomas Wh. IV. 21 (P. P. III. 16). Green keeper on golf course. Married, four children all showing deformity.

		Right and Left.
Hands.	Carpus.	Trapezium presence doubtful. Trapezoid probably not fully developed on left hand.
	Metacarpals.	Three inner bones alone present.
	Digits.	4th finger only present, very much curved, three phalanges.
Feet.	Tarsus.	Presence of middle and external cuneiforms doubtful.
	Metatarsals.	5th alone present. A small hard mass is present lying over internal cuneiform bone, possibly a rudimentary 1st metatarsal.
	Digits.	5th alone present, much curved, three phalanges.

For photographs of hands and feet: see Plate XXVII, Figs. (xiii) and (xiv).

Annie Wh. IV. 23 (P. P. III. 13). Normal.

Henry Wh. IV. 24 (P. P. III. 16). Dead, normal; three normal children, V. 34—36.

Clara Wh. IV. 26 (P. P. III. 18). Dead, normal; one normal child, V. 37.

John Wh. IV. 28 (P. P. III. 19). Normal.

We now come to *Mrs R.'s* family :

V. 1 and 2 (P. P. IV. 17, 18). Two normal sons.

V. 9, born since the first paper was written, also a normal boy; V. 10, ditto, but stillborn.

The remainder of the family are deformed. All have the little finger only on the hands, but as Pearson's Plate IX and skiagrams in the earlier paper show, this may mean very different bones missing even in the two hands of the same individual.

V. 3 = P. P. IV. 19. V. 4 = P. P. IV. 20. V. 5 = P. P. IV. 21. V. 6 = P. P. IV. 22. V. 7 = P. P. IV. 23. Accounts will be found in Pearson's paper, detailed in the cases of V. 5 and V. 7.

V. 8 = P. P. IV. 24, the "expected birth" of the first paper recorded in a footnote (p. 74) as a girl deformed in hands and feet. She differs from the other members of this branch of the family in having a single (little) finger straight on the left hand; but this single finger is crooked on the right; she has two toes (big and little) on each foot, both "clawed" round, being thus like her mother, and differing from her sister, V. 3.

Eliza Wh. V. 16 (P. P. IV. 4). Hands and feet deformed.

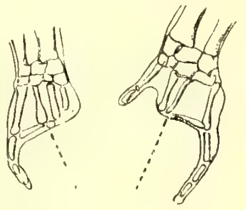
Hands	*Right*	*Left*
Carpus.	Scaphoid presence doubtful. Trapezium probably absent.	Scaphoid and trapezium probably absent.
Metacarpals.	2nd, 3rd, 4th and 5th present.	2nd (rudimentary), 3rd, 4th and 5th.
Digits.	Ring and little fingers closely united by fleshy web, double nail. Three bones in each.	Ring and little fingers closely united by a thick fleshy web, double nail. Three phalanges in each finger. Rudimentary middle finger, one small phalanx.

Feet. Same deformity in each foot.

Tarsus.	Presence of the middle and 1st cuneiform bones doubtful.
Metatarsals.	1st and 5th only present.
Digits.	1st and 5th alone present, much curved, resembling a crab's claws, normal number of bones in each digit, but joints are ankylosed.

Fred. W. V. 17 (P. P. IV. 5). Hands and feet deformed. The deformity of the hands is difficult to determine (see diagram), but probably the rigid bar of bone felt running along the free

border of the hand is due to a displaced phalanx, or to two phalanges fused and displaced*. The hands as in the majority of cases show poor muscular development; the five metacarpals are palpable on the left hand.



Rigid bars of bone.

Hands	*Right*	*Left*
Carpus.	Apparently normal.	Normal.
Metacarpals.	1st absent, 2nd rudimentary, 3rd rudimentary, but larger. 4th and 5th of normal length.	Normal but the middle three are deflected towards the thumb.
Phalanges.	In both hands a bone is felt lying between the heads of 4th and 5th metacarpals, probably a phalanx of ring finger: see diagram above.	
	5th digit shows three phalanges.	1st digit rudimentary. 5th digit, three phalanges.

The feet show the same deformity as those of *Eliza Wh.*, his sister.

For photograph of hands: see Plate XXVII, Fig. (xvii), to right.

Henry Wh. V. 18 (P. P. IV. 6). Hands and feet defective. Hands have one digit only on each; feet resemble those of his sister, *Eliza Wh.*, V. 16.

Hands, symmetrical.

Carpus.	Trapezium and trapezoid bones probably absent in both hands.
Metacarpals.	3rd metacarpal probably present but very imperfectly developed. 4th metacarpal better developed. 5th metacarpal normal.
Digits.	5th only present on each hand, much curved, and joints partially ankylosed.

For photograph of hands: see Plate XXVII, Fig. (xvii), to left.

William Wh. or *Sm.* V. 19 (P. P. IV. 7). Dead. Hands said to have one finger only present. Feet normal.

Alfred Wh. or *Sm.* V. 21 (P. P. unborn). At present in one of Dr Barnardo's homes. One finger only on each hand, the 4th, well developed. *Feet.* The outer digits alone present, though feet appear well developed save for the absence of the toes.

We are indebted to Dr C. T. Ewart of Claybury Asylum for the facts and photograph of this case.

For photograph: see Plate XXV, Fig. (vii).

William Wh. V. 30 (P. P. IV. 15). Hands and feet deformed.

Hands. Right and *left* are symmetrical.

Carpus.	Trapezium probably absent.
Metacarpals.	The inner three alone present.
Digits.	Little finger only present and has three phalanges with freely moveable joints.

* This case is of peculiar interest as showing in the *S.*-family the occurrence of the same "bridge" type or "cross-bone" type that Lewis and Embleton have illustrated by skiagrams in the *G.*-family; see *Biometrika*, Vol. VI. p. 41 and plates.

Feet. Also symmetrical.

Tarsus.	The external and middle cuneiform bones are probably absent.
Metatarsals.	5th alone present.
Digits.	A small and very much curved little toe only is present ; it probably has the normal number of bones in it.

For photograph of hands and feet : see Plate XXVI, Fig. (xi).

Thos. W. Wh. V. 31 (P. P., then unborn). Hands resemble those of his brother (V. 30) in every respect : but have more marked grooves between metacarpal heads.

Feet	*Right*	*Left*
Tarsus.	The presence of the middle and external cuneiform bones is doubtful.	Apparently normal.
Metatarsals.	5th alone present. 1st present, rudimentary.	Probably a rudimentary 1st. 5th present.
Digits.	5th alone present and curved. Small mass of fibrous tissue overlies the rudimentary 1st metatarsal.	5th normal in number of bones, but curved.

This child is only five years of age and it is difficult to say with any degree of certainty which bones are present and which absent in the case of the carpus, tarsus and the metacarpal and metatarsal bones. This is true also of the following case.

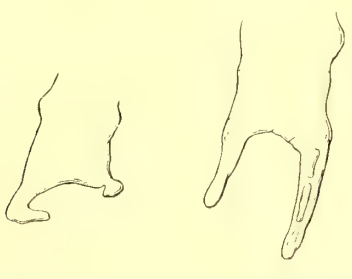
For photograph of hands and part of feet : see Plate XXVI, Fig. (xii).

R. Wh. V. 32 (P. P., then unborn). Hands and feet deformed.

The hands are symmetrical and probably have the trapezium, at least, absent. Either the inner two or three metacarpal bones are present, but are badly developed, except the innermost. One digit, the little finger, alone is present, and has the usual three bones ; it is very much curved.

The feet are asymmetrical.

	Right	*Left*
Tarsus.	Possibly normal, though the cuneiform bones are not palpable.	
Metatarsals.	1st rudimentary, 5th normal.	1st and 5th present.
Phalanges.	Three bones in 5th digit.	The first toe is long, and apparently composed of two bones. The little toe is as long as a child's middle finger, perfectly straight and without a nail. Three bones present.



For photograph of hands and feet : see Plate XXV, Fig. (viii). The photo does not give a proper idea of left foot which looks more as in diagram.

— *Wh.* V. 33 (P. P., then unborn). An infant, three months old. Hands and feet deformed, the deformity of the hands resembling those in V. 32.

The feet both show a well-developed little toe, and a mal-developed great toe.

The additional information now obtained is of much interest. We can assert:

(i) That our increased knowledge has not shown the deformity to give any closer approach to Mendelian inheritance—on the contrary the disproportionate number of abnormals has increased.

(ii) That the experience of the last five years further emphasises the importance of legislative interference with the propagation of deformity.

Pearson in 1908 laid stress on the wide variation found in the types of hand and foot, and remarked that all Lewis and Embleton's types from the *G.*-family with the exception of the "cross-bones" had been found in the *S.*-family; but he suggested that this might also be revealed if the whole family were skiagraphed. McMullan has now shown that this type actually exists.

A careful comparison of the present pedigree with the former one will show that, if additional births be put on one side, the chief changes are in the orders of birth in the earlier generations. We cannot lay much stress even now on these orders. In the class with which we have here to deal, the members frequently seem uncertain of their own ages—and are therefore ignorant of their brothers' and sisters' ages and their order of birth. When we find that they do not know their own mother's Christian name correctly, we need not be surprised that they contradict each other's statements as to the relative ages or even as to the numbers of their brothers and sisters. This especially will be the case in people of this class when the youngest is born 25 to 28 years later than the eldest. Unfortunately the Church baptismal registers appear equally chaotic; they make no pretence to identify father and mother by the same names in two successive years! On the whole there is a reasonable agreement in the two accounts. Those who have studied such families as those in the *G.** and *S.*-pedigrees of Lobster-claw, must earnestly wish for some immediate step to be taken to restrict the propagation of the physically defective; it is no less urgent than the restriction of the propagation of the mentally defective.

* We have further evidence that in the *G.*-family there has also been a considerable number of births of deformed individuals in the last five years, and a similar experience has arisen in the case of a number of pedigrees of hereditary deformity recently brought up to date by the Eugenics Laboratory staff.

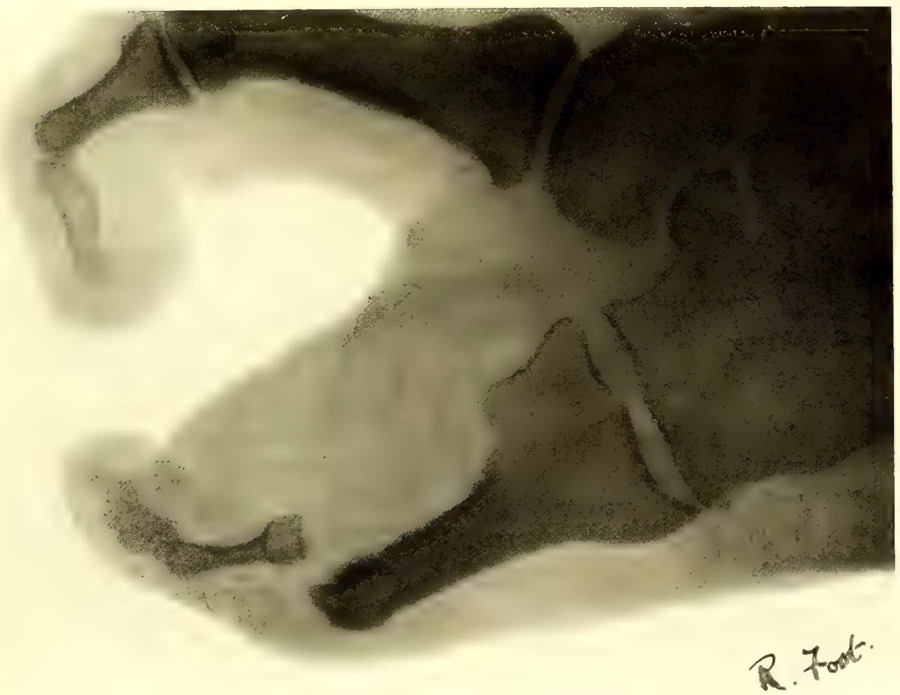


Fig. (i). *E. Sm.* Pedigree IV. 17. Right Foot. (See Plate XXV, Fig. (v).)

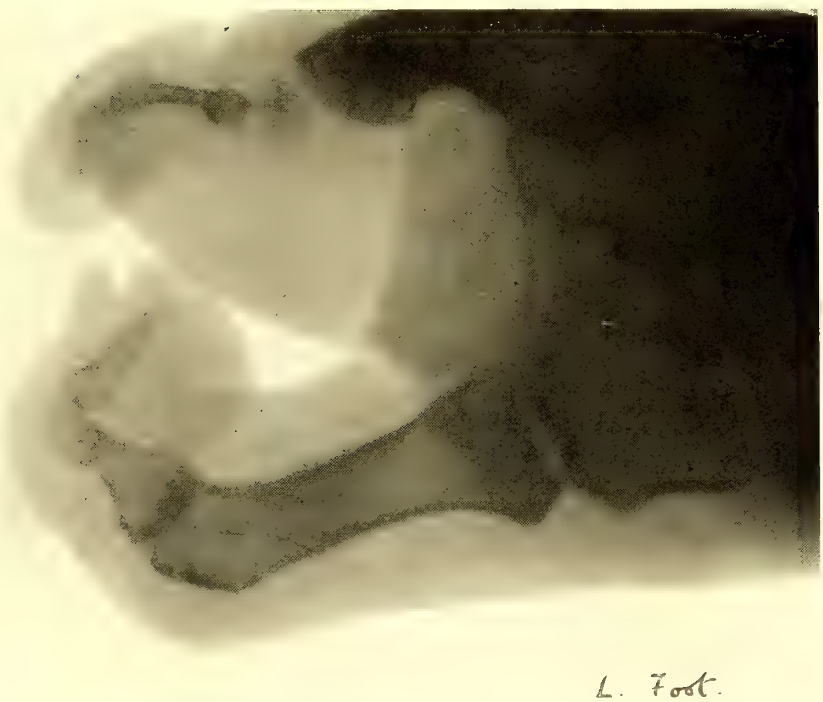


Fig. (ii). *E. Sm.* Pedigree IV. 17. Left Foot.

"LOBSTER CLAW" FAMILY.

Fig. (iii). *E. Sm.* Pedigree IV. 17. Right Hand. (See Plate XXV, Fig. (v).)

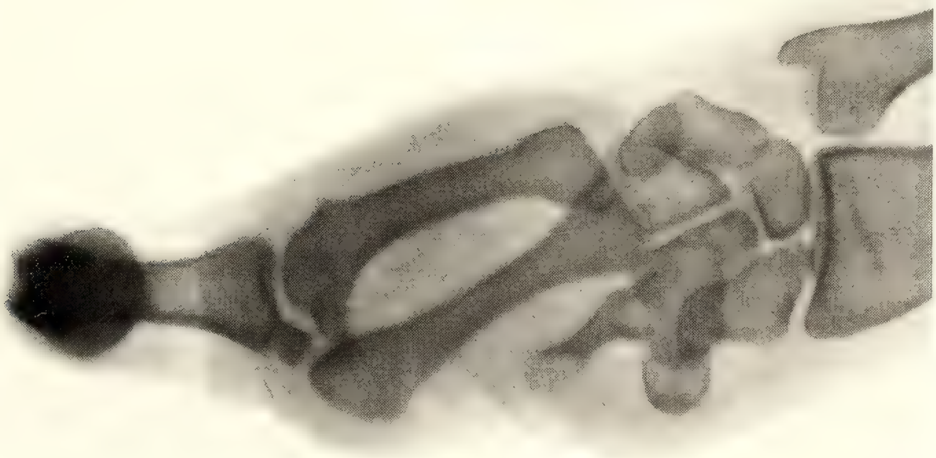


Fig. (iv). *E. Sm.* Pedigree IV. 17. Left Hand.

"LOBSTER CLAW" FAMILY.

Fig. (v). *E. Sm.* Pedigree IV. 17. Hands and Feet. (See Plates XXIII—XXIV.)

Fig. (vi). *A. E. F.* Pedigree IV. 9. Hands. (See Plate XXVII, Fig. (xviii).)

Fig. (vii). *A. Wh.* Pedigree V. 21. Right Hand, and Feet.

Fig. (viii). *R. Wh.* Pedigree V. 32. Hands and Feet.

"LOBSTER CLAW" FAMILY.

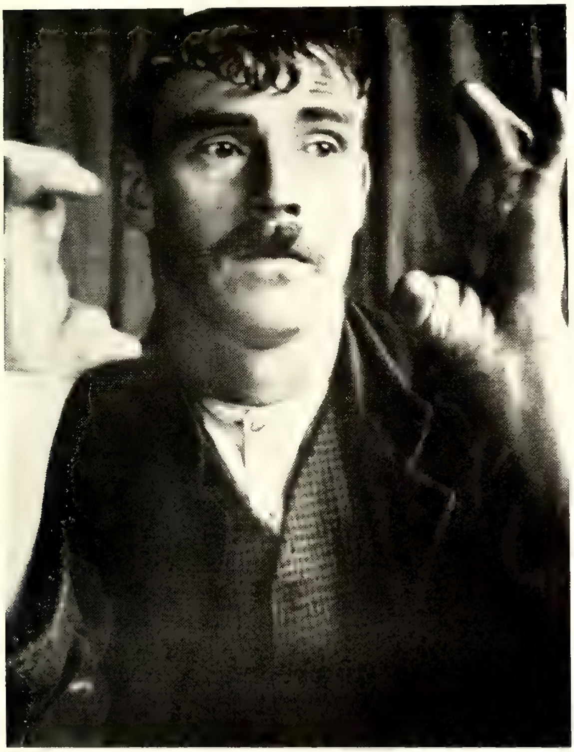


Fig. (ix). *H. S.* Pedigree IV. 11. Hands.

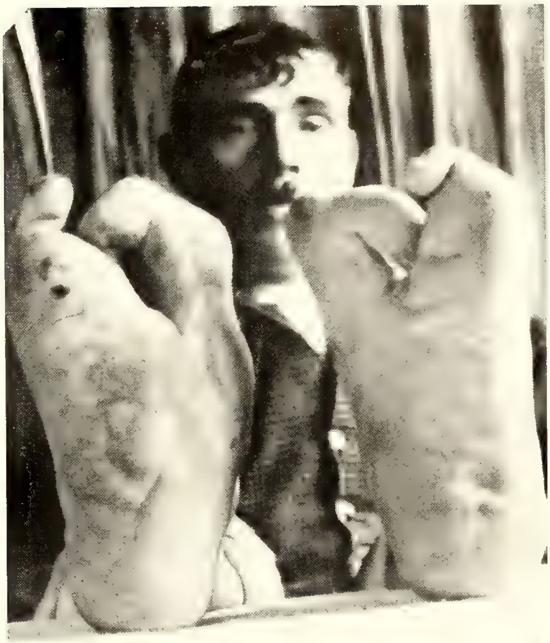


Fig. (x). *H. S.* Pedigree IV. 11. Feet.

Fig. (xi). *W. Wh.* Pedigree V. 30. Hands and Feet.

Fig. (xii). *T. W. Wh.* Pedigree V. 31. Hands and Part of Feet.

"LOBSTER CLAW" FAMILY.

Fig. (xiii). *T. Wh.* Pedigree IV. 21. Hands.

Fig. (xiv). *T. Wh.* Pedigree IV. 21. Feet.

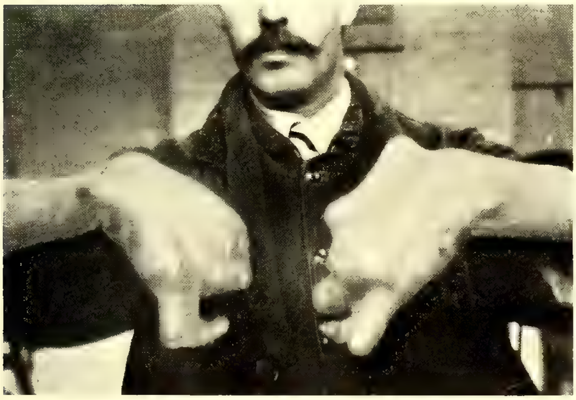


Fig. (xv). *R. S.* Pedigree IV. 7. Hands.

Fig. (xvi). *R. S.* Pedigree IV. 7. Feet.

Fig. (xvii). *H. Wh.* Pedigree V. 18. *F. W.* Pedigree V. 17. Hands.

Fig. (xviii). *A. E. F.* Pedigree IV. 9. Feet. (See Plate XXV, Fig. (vi).)

"LOBSTER CLAW" FAMILY.

A STUDY OF THE NASAL BRIDGE IN THE ANTHROPOID APES AND ITS RELATIONSHIP TO THE NASAL BRIDGE IN MAN.

By KATHLEEN V. RYLEY, Crewdson-Benington Student in Craniometry AND JULIA BELL, M.A., assisted by KARL PEARSON, F.R.S.

CONTENTS.

(1) *Introductory.*

The measurements on which the following paper is based were taken by Kathleen V. Ryley, in part at the Museum of the Royal College of Surgeons by the kindness of Professor Arthur Keith; in part at the British Museum (Natural History), where we have especially to thank Mr W. P. Pycraft; in part at the Cambridge Anatomical Museum and in the private collection of Dr Duckworth, whom we have heartily to thank for his courteous and ready aid; in part at the Oxford Museum by the kindness of Professor Bourne. Lastly we owe the possibility of measuring a few specimens to Professors Thane and Hill of University College, London, who readily placed their material at our disposal. We have further to express our indebtedness to Professor Thane and to Dr Derry for aid in a variety of other ways. Only by the great help received from all these authorities would

it have been possible to bring together even the slender numbers now dealt with, and a very full sense of our obligations to these gentlemen must here be expressed.

In a recent paper entitled "A Study of the Negro Skull" (*Biometrika*, Vol. VIII. pp. 315—320) a special study was made of the nasal bridge in the Congo and Gaboon crania and the results compared with measurements on Egyptian and English series. There resulted values so suggestive that a wider study was promised, and that promise is to some extent fulfilled in the present paper.

In the former paper it will be remembered that the chord from dacryon to dacryon (DC) was measured—the mesodacryal chord—and the minimum arc from dacryon to dacryon (DA)—the mesodacryal arc. The ratio 100 (arc − chord)/chord was termed the mesodacryal index β. If on the other hand we measure the subtense of the minimum arc (DS) and take the index 100 subtense/chord, we obtain a second mesodacryal index, which we will term here the mesodacryal index α. As α cannot be found without a special instrument, or a special construction, an attempt was made to determine the second mesodacryal index α from the first β by the hypothesis that in man the bridge of the nose may be sufficiently closely represented by a catenary, which seemed a fairly justifiable assumption in this case. The index thus indirectly found will be termed in this paper the mesodacryal index α'. Tables to determine α' from β calculated by one of us (Julia Bell) were published in the same number of *Biometrika*, pp. 338—9, as the paper on the negro skull. An extension of these Tables due to H. E. Soper, rendered necessary when we deal with apes, will be found in the footnote on p. 401 of the present paper. Meanwhile our attention was drawn by Professor G. D. Thane to Mérejkowsky's pioneer work of 1882 on the nasal bones. He measured the shortest horizontal chord from nasal suture to nasal suture (SC) and the minimum subtense (SS) and took the index 100 SS/SC, which we have ventured to term the *Simotic Index*, S. There is little doubt that this index is of a more simple anatomic character, as being only concerned with the nasal bones, than the mesodacryal indices. But it is by no means so marked a physiognomic feature in the living as the mesodacryal index. It has further to be obtained by aid of a special instrument—in our case a modified Mérejkowsky instrument,—which may be termed the simometer. This simometer may also be used to test the relationship between observed α and calculated α'.

In the present investigation α, β, S have been measured and α' calculated for Gorilla, Chimpanzee, Gibbon and Orang. But in the case of the latter ape it was found possible to measure the simotic index in only a few cases,—the index taking frequently the form of the ratio of two almost vanishing quantities*. In

* There were three cases in which the simotic chord was recorded among the Orang-utans. In two males it was 5·7 and 5·8 mm. In both these cases the subtense was zero, for the bones were practically flat; thus the simotic index was zero. In the third case, a female, the simotic chord was 7·0 mm. and the subtense 1·2 mm. and therefore the simotic index was 17·1. There were a few further cases in which the chord only was measurable. Thus on the basis of such isolated cases we should have to place the Orang-utan below the Chimpanzee, and both of course below the Negroes.

the Orang the nasal bones are fused into a narrow flat strip from 5 to 6 mm. wide, sometimes there is only a thread at the nasal bridge or the nasals are entirely absent. In some cases the maxillary bones stand out forming a ledge which overhangs the region of the dacryon and extends beyond it, thus making it very difficult to measure with Mérejkowsky's instrument. Initially we excluded entirely from our measurements the Gibbon, but notwithstanding the difficulties of including this ape, it afterwards appeared to us that it might be worth while to indicate the results in rough numbers. The difficulties partly arise from the fact that a large proportion of the Gibbon crania available are without any statement as to sex and we did not feel that an attempt at sexing would be really profitable, as the total numbers were not large enough to give double series. Further we are compelled to distinguish between *Hylobates* and *Symphalangus*, and even of the former with the sparse data as to locus usually provided, it is impossible to assert that we are dealing with anything like a homogeneous race. Thus we found ourselves ultimately with only 25 cases—♂ + ♀ + ☿ —*Hylobates*—indeed only 16 for the simotic characters, and 9 cases—♂ + ♀ + ☿ —*Symphalangus*—with merely six for the simotic characters. Our results in this case can therefore only be rough indications of the racial characters of the nose of these apes. Our Museums have yet to learn that for comparative study we want at least 50 to 100 crania of each sex drawn from definite districts. To compare with those of the apes the corresponding indices were measured in a number of human races selected in order to contrast the values, if possible, with those of the anthropoid apes of the same continent. A few other human races are taken to indicate general relationships. The Hindoo, Egyptian and English crania used are at University College either in the Biometric Laboratory or in the Department of Anatomy; the Asiatic skulls and the Congo skulls are at the Royal College of Surgeons; the Gaboon skulls at the British Museum of Natural History. The nasal features of the latter as well as of the English and Egyptian crania had already been discussed in Dr Crewdson-Benington's paper referred to above.

(2) *Reliability of Measurements.*

The first point to be considered is the degree of reliability in the measurements made, and here we must again emphasise the difficulty of accuracy. Not only do the quantities measured differ by their extreme smallness from the customary cranial measurements, but the points for measurement and the instrument for measurement are by no means all that could be desired. When a small quantity has to be measured, an error of reading which would have no importance in the case of either head breadth or head length, the cephalic index being in question, becomes of very great significance in the case of the simotic or even the mesodacryal indices. Further in the handling of Mérejkowsky's instrument, there is much play for personal equation, and only after long practise will the recorder settle down into a process, which repeats within reasonable limits the values of the indices previously determined. In the next place for our present purposes only small series of crania were as a rule available, e.g. 20 to 40 of either

sex of each ape, 10 to 40 of the various Asiatic races. Considering the smallness of these numbers, it did not appear needful, in view of the labour involved, to measure more than 50 of each sex of the English (Whitechapel) and Egyptian (XXVI—XXX dynasties) crania of which far larger numbers were at hand.

Further in such small series the question of sexing becomes of considerable influence. With a large series of crania of one type, we sex on our knowledge of that series itself. But in short series one is apt to sex by *a priori* impressions of sex in other races, because we have not before us wide enough material to appreciate sex differences within the race itself.

We propose to illustrate first the order of the differences which may arise from these sources, and we will deal first with sexing. Owing to the courtesy of Dr Derry we were able to test this point on 35 Nubian crania in his possession. They were independently sexed by Professor G. D. Thane and by Dr Derry himself*. The nasal measurements were taken and the mean values of the indices calculated. The following values were obtained, where we have added for comparative purposes those of other negro series:

TABLE I.

	m_a		$m_{a'}$		m_β		m_S	
	♂	♀	♂	♀	♂	♀	♂	♀
Nubians: Professor Thane ...	48·0	40·9	49·7	43·6	49·6	40·0	34·5	23·4
,, Dr Derry	48·9	39·7	51·0	42·1	51·6	37·7	33·7	23·5
Gaboon: 1864	42·9	44·0	43·6	44·9	40·0	42·0	30·9	27·7
,, 1880	42·9	44·1	44·8	45·1	41·9	42·3	29·0	22·9
Congo	39·1	37·2	40·5	38·6	35·5	32·6	25·6	25·7
"Nègres": Mérejkowsky ...	—	—	—	—	—	—	25·6	

It will be seen that for short series the personal equation of sexing makes 1 to 2 points difference in the indices, but that in two series of the same race—Gaboon, 1864 and 1880,—differences of this order may occur at least in the simotic index. If we compare Mérejkowsky's value for the simotic index with ours, we see that there would be excellent agreement if his somewhat vague "Nègres" were from the Congo district. But we lay no stress on this agreement because we believe that Mérejkowsky tends almost invariably to get a higher simotic index than we do. This is illustrated in Table II.

Of course there are but few races common to both series and in Mérejkowsky's series as in ours the numbers dealt with are miserably small, but still cases where

* Before discussion there were four crania in a total of 35 on which there was difference of opinion, say 11 per cent., and we have used the original determinations, putting as usual ?♂ and ?♀ with the ♂'s and ♀'s respectively.

we can compare and the general appearance of the Table are both against the view that the differences are wholly racial. We believe them to be largely due to personal equation in taking a delicate measurement with a not very reliable instrument, and consider that without determination of personal equations by direct comparison of measurements on the same individual crania, it is not possible at present to compare simotic indices found by different craniologists.

TABLE II.

Values of the Simotic Index.

Race	Mérejkowsky		Race	Ryley	
	♂	♀		♂	♀
Negroes	25·6	—	Congo	25·6	25·7
Malays	31·3	—	Moluccas	26·5	—
New Caledonians ...	38·5	30·6	Gaboon, 1880 ...	29·0	22·9
Mongols	40·5	—	Sumatra	29·2	—
Melanesians... ...	41·9	35·2	Celebes	29·7	—
Marquesas Island ...	43·9	34·0	Borneo	30·2	34·7
Maoris	47·9	43·3	Philippines	30·4	29·0
American Indians ...	48·0	46·2	Gaboon, 1864 ...	30·9	27·7
New Hebrides ...	49·1	44·5	Malay	33·7	—
Hindoos	51·1	42·5	Java and Madura ...	33·9	32·7
Auvergnats	51·8	45·0	Nubians	34·1	23·4
Gypsies	53·8	—	Ainos	43·4	—
Tahitians	54·3	—	Veddahs	43·6	36·9
French	54·8	47·1	Egyptians	44·4	36·9
Dutch	58·6	55·2	Hindoos	44·7	39·1
French Cymry ...	59·6	50·3	English	50·8	46·6

In two points, however, we agree fairly well with Mérejkowsky, i.e. in the general order of races and in the fact that the simotic index is a marked sexual character. In all cases—22 in number—except the Congo and Borneo crania, the female has a lower simotic index than the male and this is true for the apes, where we can determine it:

	♂	♀
Chimpanzee	21·4	18·5
Gorilla	56·8	55·1

The two exceptions are the Congo crania, where the two sexes are practically alike, and the crania from Borneo, where the female has the higher simotic index, but the probable errors being introduced, i.e. Borneo: ♂'s $30·2 \pm 1·2$, ♀'s $34·7 \pm 2·1$, show us that even here the difference is quite possibly non-significant. Thus we conclude that in the female the nasal bones are flatter than in the male, and this secondary sexual character may be used—as far as it is ocularly appreciable—as a help in sexing. It may be remarked—in order to avoid circular reasoning—that

with the possible exception of Dr Derry's classifications, the judgment in sexing has not been influenced by nasal characters*.

While it appears to us that the difference of sexing will not be greater than that of two small samples of the same race and less probably than the difference of measurement of two different craniologists, we have still to answer the important question of how far the same craniologist will repeat results in the case of these delicate nasal measurements. We are able to make some interesting comparisons from this standpoint. K. V. Ryley made her first measurements of the nose on the Whitechapel, Egyptian, Congo and Gaboon series in the summer of 1911 for use in the paper on the negro skull (*Biometrika,* Vol. VIII. p. 316). She was then new to these particular measurements, and before the second series were made she had nearly a year's practice on a great variety of races. In the English Whitechapel series 13 crania measured in one series (1911) are not in the other series (1912), that is to say one quarter of the crania are not the same in the two series. These 13 crania of the first series were originally Museum specimens, i.e. skulls having some abnormality (other than nasal), or remarkable feature, and placed in the Anatomical Museum for exhibition. It was thought better to exclude these individuals in the second series and replace them by more normal examples. Thus the divergencies between the A (1911) and B (1912) English series are not solely due to personal equation. The Congo and Gaboon 1864 series are practically the same crania for A and B as nearly every skull which was measurable was measured on each occasion.

The two Egyptian series are not the same crania, for they consisted of the 100 crania of the series of 1800 skulls, which were being otherwise measured at the time when these nasal measurements were taken. Thus we may conclude that the Congo and Gaboon 1864 series represent changes due to the same measurer repeating operations at the beginning and end of a period during which experience was growing and methods being standardised; the Egyptian series represent this variation in the measurer together with the difference between random samples; and the English series represent not only the variation due to the measurer, but further the effect of modifying by 25 °/。 the actual sample

* Dr Derry does give weight to nasal features in sexing, but if we examine the results in Table I from this standpoint, we have

Difference of male and female, i.e. ♂ − ♀.

	m_a	$m_{a'}$	m_β	m_S
Professor Thane	7·1	6·1	9·6	11·1
Dr Derry ...	9·2	8·9	13·9	10·2

It will be clear from these results that Dr Derry used the mesodacryal indices, i.e. the entire bridge of the nose, rather than the nasal bones in his sexual appreciation; for he has a less sexual difference in the simotic index than Professor Thane, while he has greater values in the mesodacryal indices. As a matter of fact the sexual differences in the Nubians seem to us of such an exaggerated character, that we venture to suggest that the females may have been captives and thus may be more definitely negresses, with whom their indices accord, than the males were negroes; the males being possibly a conquering race.

measured. Finally to test the extent to which K. V. Ryley's methods were finally standardised, a third measurement of the male Whitechapel crania was made four months later than *B*, namely *C* in December, 1912. In the case of *C* the same 50 crania exactly were remeasured*. There now resulted a very close accordance between all the measurements absolute and indicial. It will be noticed on examination of the Table that:

(i) The error in the value of the simotic index due to variation of personal equation never exceeds a point.

TABLE III.

Comparison of Measurements on the Nasal Bridge made at intervals and on different samples.

Series	Meso-dacryal Chord (*DC*)	Meso-dacryal arc (*DA*)	Simotic Chord (*SC*)	Simotic Subtense (*SS*)	Meso-dacryal Index β	Meso-dacryal Index α' (from β)	Simotic Index (*S*)	Remarks
Gaboon, 1864, *A* ♂	24·5 ± ·2	33·5 ± ·3	9·16 ± ·26	2·87 ± ·11	42·7 ± ·9	45·4 ± ·6	30·8 ± ·9	Same crania
" " *B* ♂	23·2 ± ·2	32·4 ± ·3	9·43 ± ·26	2·88 ± ·11	40·0 ± ·8	43·6 ± ·5	30·9 ± ·9	
" " *A* ♀	23·1 ± ·3	32·3 ± ·4	9·55 ± ·28	2·68 ± ·10	44·8 ± 1·3	46·7 ± ·8	28·7 ± 1·0	Same crania
" " *B* ♀	22·5 ± ·3	31·8 ± ·4	9·88 ± ·29	2·68 ± ·11	42·0 ± 1·3	44·9 ± ·8	27·7 ± 1·0	
Congo, *A* ♂ ...	24·4 ± ·3	32·8 ± ·4	9·44 ± ·21	2·38 ± ·08	34·5 ± ·9	39·6 ± ·6	25·6 ± ·8	Same crania
" *B* ♂ ...	23·9 ± ·3	32·4 ± ·4	9·80 ± ·21	2·48 ± ·09	35·5 ± ·8	40·5 ± ·6	25·6 ± ·7	
" *A* ♀ ...	23·5 ± ·3	32·2 ± ·4	9·46 ± ·33	2·31 ± ·11	36·8 ± 1·0	41·5 ± ·7	25·2 ± 1·3	Same crania
" *B* ♀ ...	23·3 ± ·3	30·8 ± ·3	9·72 ± ·35	2·41 ± ·12	32·6 ± ·8	38·6 ± ·6	25·7 ± 1·4	
Egyptians, *A* ♂ ...	21·5 ± ·2	34·5 ± ·4	10·58 ± ·18	4·67 ± ·10	61·3 ± 1·5	57·0 ± ·9	44·8 ± ·8	Different samples
" *B* ♂ ...	21·5 ± ·2	35·3 ± ·2	10·42 ± ·17	4·63 ± ·10	65·0 ± 1·3	59·4 ± ·8	44·4 ± ·7	
" *A* ♀ ...	20·4 ± ·2	32·4 ± ·2	10·76 ± ·17	4·11 ± ·09	59·3 ± ·9	55·9 ± ·6	38·6 ± ·7	Different samples
" *B* ♀ ...	20·6 ± ·3	32·0 ± ·3	10·84 ± ·16	3·97 ± ·09	56·0 ± 1·1	53·9 ± ·7	36·9 ± ·8	
English, *A* ♂ ...	22·6 ± ·2	35·1 ± ·3	9·52 ± ·17	4·74 ± ·11	56·2 ± 1·3	53·9 ± ·8	50·7 ± 1·2	25 °/ₒ different crania
" *B* ♂ ...	21·5 ± ·2	34·8 ± ·3	9·38 ± ·17	4·68 ± ·10	62·7 ± 1·2	58·0 ± ·7	50·8 ± 1·2	
" *A* ♀ ...	20·9 ± ·2	31·9 ± ·3	8·98 ± ·18	4·14 ± ·07	53·4 ± 1·2	52·2 ± ·8	47·4 ± ·9	25 °/ₒ different crania
" *B* ♀ ...	20·6 ± ·2	31·5 ± ·3	9·10 ± ·17	4·16 ± ·07	53·7 ± 1·0	52·5 ± ·6	46·6 ± ·8	

TABLE III *bis*.

English Crania, 50 Males.

Series	Meso-dacryal Chord (*DC*)	Meso-dacryal arc (*DA*)	Meso-dacryal Subtense (*DS*)	Simotic Chord (*SC*)	Simotic Subtense (*SS*)	Meso-dacryal Index β	Meso-dacryal Index α'	Meso-dacryal Index α	Simotic Index *S*
B ...	21·5 ± ·2	34·8 ± ·3	12·1 ± ·1	9·38 ± ·17	4·68 ± ·10	62·7 ± 1·2	58·0 ± ·7	56·8 ± ·7	50·8 ± 1·2
C ...	21·4 ± ·2	34·6 ± ·3	12·1 ± ·1	9·43 ± ·17	4·72 ± ·10	62·7 ± 1·2	58·1 ± ·7	57·4 ± ·7	51·2 ± 1·2

A are measurements taken at start of investigation, *B* at conclusion of investigation after practise and standardisation, *B* only used in this memoir. In the Table III *bis* both series, *B* and *C*, were taken on the same crania at the end of the investigation with an interval of four months.

* K. V. Ryley, it is needless to say, had not her earlier values before her when making her remeasurements.

That is, it is of the order of the probable error. It may amount to 1·7, which is only about twice the probable error when we combine the effect of random sampling with personal error.

(ii) The differences in the mesodacryal index β can amount for personal variation to 4·2 (Congo ♀'s), which is as large a value as we obtain on the Egyptian series for combined personal error and random sampling, and even to 6·6 points in the Whitechapel with change of sample and personal variation, which is five times the probable error. Corresponding differences will of course exist in α' deduced from β. We believe that this difference in the early and late values of β is due to greater standardisation of method in determining the dacryon with a rather blunt pointed instrument and to greater mastery of the difficulties arising from damaged crania and in the case of the apes from partially obliterated sutures.

(iii) In the absolute measurements mesodacyral chord (DC), mesodacryal subtense (DS), mesodacyral arc (DA), simotic chord (SC), simotic subtense (SS) there is a maximum error in the means of about 4·5 °/$_{\circ}$, but this is rare, and reduces to insignificance with practise in the measurements*.

(iv) A comparison of B and C shows that in all cases after continuous practise and standardisation of method, the variations in absolute measurements or in indices are less than the probable errors of random sampling.

Of course the whole of the above remarks apply to a comparison of *mean* values and not to measurements on an individual cranium. We should conclude as follows:

(a) The measurements on the nasal bridge need great care, and the same measurer may on repetition differ considerably from a previous determination on the same skull.

(b) But with long practise and careful standardisation of method the racial values obtained will repeat themselves, if the series be measured afresh, and values obtained by the same recorder are quite intercomparable.

(c) We are doubtful—especially having regard to Mérejkowsky's results—whether racial values obtained by different craniologists, who are fresh to this side of cranial measurement, or who have not tested their personal equation on the same series, will be directly comparable.

(d) Notwithstanding these difficulties, which are really only an intensification of difficulties which occur in all branches of craniometry, we believe these measurements on the nasal bridge to be really valuable and think that they ought to occupy more attention. They are of course only a phase of what may reach greater importance in the future: namely of what we may term *micrometric* as distinguished from the usual or *macrometric* measurements of the skull.

* The earlier series of measurements tended rather to exaggerate the mesodacryal chord and arc.

There are bones other than those of the nasal bridge which might give racially significant micrometric measurements, although specially designed apparatus, or an adaption of existing physical instruments as yet unused in craniometry, would have to be employed in their determination.

(3) *On the observed* (α) *and calculated* (α') *values of the subtense mesodacryal index.*

In the discussion on the negro skull (*loc. cit.* pp. 297, 316) it was pointed out that our original measurements of the mesodacryal chord and arc leading to the index β did not enable us to determine the subtense directly, or indirectly, until some assumption was made as to the shape of the bridge of the nose. We were not at the time these measurements were suggested to Dr Crewdson-Benington aware of Mérejkowsky's instrument, which is as available for finding the mesodacryal subtense as for its designed purpose the determination of the simotic subtense. Accordingly after some consideration the catenary was selected as a fairly suitable curve for the nasal bridge and by means of it the value of the subtense mesodacryal index (α') was deduced from β. In the present investigation α has been found directly and it is of much interest to test the legitimacy of the catenary hypothesis.

TABLE IV. *Differences* $\alpha' - \alpha$ *of Calculated* (α') *and Observed* (α) *Values of Mesodacryal Subtense Index.*

	♂	♀	Average
Gorilla	0·21	0·07	0·14
Ainos	0·36	—	0·36
Negroes, Gaboon (1864 + 1880)	1·01	0·88	0·94
Negroes, Congo	1·45	1·37	1·41
English	1·19	−0·47	0·36
Egyptians	1·53	1·12	1·32
Chimpanzees	1·64	1·41	1·52
Philippine Islands	1·10	2·16	1·63
Moluccas	1·63	—	1·63
Hindoos	1·79	3·12	2·45
Celebes	2·41	—	2·41
Borneo	2·76	3·86	3·31
Javanese	2·20	2·80	2·50
Sumatra	3·73	—	3·73
Veddahs	4·58	4·99	4·78
Orang-utans	17·33	20·23	18·78
Gibbons	♂ + ♀ + ♂ *Hylobates* 18·60		18·60
	♂ + ♀ + ♂ *Symphalangus* 14·22		14·22

We see, to judge by this Table, that the hypothesis of the catenary suits the Gorilla and the Aino nose very closely; it is fairly good for Negro, English and

Egyptian, roughly we must subtract one point from the calculated to get the actual value of the index; for the Chimpanzee, the Hindoos and the inhabitants of the Philippine and Molucca Islands two points must be subtracted; for those of the Celebes, Borneo, Java and Sumatra three points, for the Veddahs five points, and for the Orang-utan and Gibbon 14 to 20! Thus as we pass from Africa and the region of the Gorilla through Europe to Asia and to the East Indian Islands and the region of the Orang-utan we get roughly a continuous divergence from the catenary hypothesis, which fits the Gorilla, to a markedly non-catenary relation as in the case of the Orang*. While the Aino and Veddah have equal flatness of the nasal bones as measured by the simotic index, they stand at opposite ends of the scale with regard to a catenary hypothesis of the whole bridge of the nose. No attempt, however, is made to deduce any racial order from Table IV although it is not without suggestiveness.

Another method of approaching the degree of relationship of α' to α is to consider the correlation coefficient between them, $r_{\alpha\alpha'}$. A glance at Table XX shows that it is very high. We have, averaging results:

TABLE V. *Correlation* $r_{\alpha\alpha'}$.

Chimpanzees and Orang-utans	·754
Gibbons (Hylobates)	·815
East Asiatic Group	·878
Gorillas	·906
Negro Group	·923
Hindoos	·931
Egyptians and English	·964
Veddahs	·965

Finally in order to calculate the value of α from α' for a given individual of one of the races dealt with, we append the regression lines, which provide the most probable value of α for a given α'.

TABLE VI.

Male		Female	
Orang-utans	$a = \cdot 6697a' + 12\cdot 99$	*Orang-utans*	$a = \cdot 5053a' + 28\cdot 98$
Chimpanzees	$a = \cdot 6629a' + 8\cdot 66$	*Chimpanzees*	$a = \cdot 6096a' + 10\cdot 50$
Veddahs	$a = \cdot 8324a' + 5\cdot 75$	*Gorillas*	$a = \cdot 8404a' + 7\cdot 11$
Congo	$a = \cdot 8431a' + 4\cdot 91$	Egyptians	$a = \cdot 8877a' + 4\cdot 93$
Egyptians	$a = \cdot 9359a' + 2\cdot 25$	Gaboon, 1880	$a = \cdot 8842a' + 4\cdot 28$
Philippine	$a = \cdot 9448a' + 1\cdot 25$	Gaboon, 1864	$a = \cdot 8931a' + 3\cdot 94$
Javanese	$a = \cdot 9359a' + \cdot 82$	English	$a = 1\cdot 0037a' + \cdot 27$
Gaboon, 1864	$a = \cdot 9782a' + \cdot 23$	Congo	$a = \cdot 9739a' - \cdot 36$
English	$a = \cdot 9756a' + \cdot 22$	"Mixed" Malays	$a = 1\cdot 0420a' - 5\cdot 76$
Hindoos	$a = \cdot 9665a' + \cdot 11$	Veddahs	$a = 1\cdot 1203a' - 12\cdot 71$
Borneo	$a = \cdot 9645a' - 1\cdot 16$	[*Gibbons, Hylobates* ♂ + ♀ + ⚲	$a = \cdot 7083a' - 2\cdot 41$]
Gaboon, 1880	$a = 1\cdot 0045a' - 2\cdot 09$		
Gorillas	$a = 1\cdot 0750a' - 3\cdot 29$		

* The reader will observe (see contours, pp. 428—9) that it is not proper to speak of a catenary *form* of the nose. The English and Gorilla noses are not at all catenary in contour, but the chord and subtense relation is on the average clearly that of the catenary chord and subtense.

It will be seen that the catenary assumption as to the bridge of the nose, notwithstanding the relatively high correlations of α and α', is not very satisfactory *. If it were, the best fitting or regression line would be simply $\alpha = \alpha'$. Thus the regression coefficient should be unity and the constant term zero. This is roughly true for English females, less true for English and Hindoo males, and approximately the case in Congo females and Gaboon (1864) males. For the anthropoids the deviations are very considerable, and this is also true for the Veddahs. The catenary form of the nose appears to be an impression largely based on familiarity with the higher racial types. Thus the mesodacryal subtense index α can only be deduced from the mesodacryal arcual index β for *individuals* of the higher races. At the same time Table IV shows that the *racial* value α' of α as deduced from β is very little in excess of the true value for a number of races, and we conclude that the hypothesis has greater value for interracial than for intraracial comparisons†.

(4) *On the Simotic and Mesodacryal Indices.*

If we are going to use these indices as a racial scale, it becomes of some importance to determine which provides the better classification. Unfortunately the simotic index cannot be determined for the Orang-utan. But an examination of Table IX shows that the Chimpanzee is placed at one end, the Gorilla at

* It serves excellently, however, as a control of individual measurements, and when α' is not approximately α it is always worth reconsidering the measurements involved.

† In order to obtain the values of α' from those of β for the catenary hypothesis in the case of the Chimpanzee and Orang-utan, the table published in *Biometrika*, Vol. VIII. pp. 338—9, had to be somewhat extended. The values in Table VII below were obtained.

TABLE VII.

Supplementary Tables for Subtense Mesodacryal Index α as calculated from the arcual value β on the Catenary Hypothesis.

(A) Values of α.

β	·0	·1	·2	·3	·4	·5	·6	·7	·8	·9
6	15·3	15·4	15·6	15·7	15·8	16·0	16·1	16·2	16·3	16·5
7	16·6	16·7	16·8	17·0	17·1	17·2	17·3	17·4	17·6	17·7
8	17·8	17·9	18·0	18·1	18·3	18·4	18·5	18·6	18·7	18·8
9	18·9	19·1	19·2	19·3	19·4	19·5	19·6	19·7	19·8	19·9
10	20·0	20·1	20·2	20·3	20·4	20·6	20·7	20·8	20·9	21·0
11	21·1	21·2	21·3	21·4	21·5	21·6	21·7	21·8	21·9	22·0
12	22·1	22·2	22·3	22·4	22·5	22·6	22·7	22·8	22·9	23·0

the other end of the simotic scale*, while in the case of the mesodacryal indices the Chimpanzee and Gorilla although widely separated from each other are both placed at the low end of the scale next the Negroes, while the Orang-utan appears next the Veddahs at the top of the scale. This appears an arrangement at least more in keeping with geographical relations, although we make no assertion that similarity in nasal form between Veddah and Orang and again between Negro and Gorilla or Chimpanzee is due either to closer descent relationship, or to a common environment compelling species dwelling under it to follow similar nasal developments†. As far as the mesodacryal indices are concerned, English, Egyptian and Hindoo are closer to the Veddah than they are to the Negro, and may in nasal form be considered closer to the Orang than to

TABLE VII. (*continued*).

(B)

β	α	β	α	β	α	β	α	β	α	β	α
101	80·2	*126*	94·0	*151*	107·5	*176*	120·8	*201*	133·9	*226*	147·0
102	80·8	*127*	94·5	*152*	108·0	*177*	121·3	*202*	134·5	*227*	147·5
103	81·3	*128*	95·1	*153*	108·5	*178*	121·8	*203*	135·0	*228*	148·0
104	81·9	*129*	95·6	*154*	109·1	*179*	122·4	*204*	135·5	*229*	148·6
105	82·4	*130*	96·2	*155*	109·6	*180*	122·9	*205*	136·0	*230*	149·1
106	83·0	*131*	96·7	*156*	110·2	*181*	123·4	*206*	136·6	*231*	149·6
107	83·5	*132*	97·2	*157*	110·7	*182*	124·0	*207*	137·1	*232*	150·1
108	84·1	*133*	97·8	*158*	111·2	*183*	124·5	*208*	137·6	*233*	150·6
109	84·6	*134*	98·3	*159*	111·8	*184*	125·0	*209*	138·1	*234*	151·2
110	85·2	*135*	98·9	*160*	112·3	*185*	125·5	*210*	138·7	*235*	151·7
111	85·8	*136*	99·4	*161*	112·8	*186*	126·1	*211*	139·2	*236*	152·2
112	86·3	*137*	99·9	*162*	113·4	*187*	126·6	*212*	139·7	*237*	152·7
113	86·9	*138*	100·5	*163*	113·9	*188*	127·1	*213*	140·2	*238*	153·2
114	87·4	*139*	101·0	*164*	114·4	*189*	127·7	*214*	140·8	*239*	153·8
115	88·0	*140*	101·6	*165*	114·9	*190*	128·2	*215*	141·3	*240*	154·3
116	88·5	*141*	102·1	*166*	115·5	*191*	128·7	*216*	141·8	*241*	154·8
117	89·1	*142*	102·7	*167*	116·0	*192*	129·2	*217*	142·3	*242*	155·3
118	89·6	*143*	103·2	*168*	116·5	*193*	129·8	*218*	142·8	*243*	155·8
119	90·1	*144*	103·7	*169*	117·1	*194*	130·3	*219*	143·4	*244*	156·3
120	90·7	*145*	104·3	*170*	117·6	*195*	130·8	*220*	143·9	*245*	156·9
121	91·2	*146*	104·8	*171*	118·1	*196*	131·3	*221*	144·4	*246*	157·4
122	91·8	*147*	105·3	*172*	118·7	*197*	131·8	*222*	144·9	*247*	157·9
123	92·3	*148*	105·9	*173*	119·2	*198*	132·4	*223*	145·4	*248*	158·4
124	92·9	*149*	106·4	*174*	119·7	*199*	132·9	*224*	146·0	*249*	159·0
125	93·4	*150*	106·9	*175*	120·2	*200*	133·4	*225*	146·5	*250*	159·5

* There is slender evidence also to place the Orang-utan out of position below the Chimpanzee and Negroes in the matter of this index.

† Dr Hans Friedenthal (*Beiträge zur Naturgeschichte des Menschen*, Th. v. "Sonderform der menschlichen Leibsbildung," Jena, 1910, S. 61) considers the nasal form determined by environmental conditions: "Die anthropoiden Affen als Baumbewohner bedurften keiner Vervolkommung der Atemröhre während der Reduktion des Riechorgans, wie die bodenbewohnenden Menschen und Paviane. Die Gorilla, mehr dem Bodenleben angepasst als die übrigen Anthropoiden, ja, mehr als viele Ostaffen und Westaffen, zeigt eine ausgebildetere äussere Nase als Schimpanse, Orang und Hylobates."

TABLE VIII. *Man and the Anthropoid Apes. Means of Nasal Bridge Measurements.*

		Mesodacryal Chord	Mesodacryal Arc	Mesodacryal Subtense	Simotic Chord	Simotic Subtense	Index α	Index α'	Index β	Index S
Males	Congo	50 23·93 ± ·26	50 32·39 ± ·38	50 9·32 ± ·15	50 9·80 ± ·21	50 2·48 ± ·09	50 39·09 ± ·54	50 40·54 ± ·57	50 35·48 ± ·83	50 25·58 ± 1·73
	Gaboon, 1864	48 23·16 ± ·18	48 32·38 ± ·28	48 9·91 ± ·13	47 9·43 ± ·26	47 2·88 ± ·11	48 42·91 ± ·55	48 43·63 ± ·51	48 39·96 ± ·75	47 30·90 ± ·87
	Gaboon, 1880	16 23·16 ± ·46	16 32·64 ± ·40	16 9·78 ± ·19	16 8·30 ± ·40	16 2·30 ± ·11	16 42·93 ± 1·29	16 44·82 ± 1·25	16 41·93 ± 1·91	16 29·01 ± 1·49
	Philippine Islands ...	17 20·31 ± ·29	17 28·12 ± ·35	17 8·42 ± ·24	17 8·42 ± ·31	17 2·55 ± ·13	17 41·55 ± 1·08	17 42·65 ± ·83	17 38·48 ± 1·26	17 30·40 ± 1·09
	Moluccas	8 21·41	8 30·19	8 9·15	8 9·36	8 2·48	8 42·86	8 44·49	8 41·15	8 26·53
	Borneo	26 21·25 ± ·32	26 30·06 ± ·32	26 8·86 ± ·14	26 7·93 ± ·24	26 2·33 ± ·10	26 42·24 ± 1·06	26 45·00 ± 1·04	26 42·19 ± 1·65	26 30·15 ± 1·24
	Malays	14 21·12 ± ·48	14 29·96 ± ·64	14 9·23 ± ·22	14 7·76 ± ·35	14 2·62 ± ·17	14 44·16 ± 1·14	14 45·25 ± 1·24	14 42·50 ± 1·90	14 33·71 ± 1·41
	Celebes	9 21·31	9 30·78	9 9·40	9 8·54	9 2·58	9 44·52	9 46·93	9 45·12	9 29·74
	Sumatra	9 22·03	9 32·56	9 9·86	8 10·10	8 2·88	9 44·73	9 48·46	9 47·44	8 29·18
	Java and Madura ...	39 21·63 ± ·23	39 31·36 ± ·33	39 9·63 ± ·18	39 8·83 ± ·22	39 2·99 ± ·12	39 44·95 ± ·94	39 47·15 ± ·95	39 45·60 ± 1·48	39 33·88 ± 1·21
	Nubians*	17 24·21	17 36·49	17 11·72	17 11·45	17 3·74	17 48·47	17 50·37	17 50·60	17 34·07
	Ainos	6 20·88	6 31·42	6 10·42	6 8·80	6 3·80	6 49·92	6 50·28	6 50·45	6 43·37
	Hindoos	10 19·47 ± ·44	10 31·20 ± ·68	10 10·66 ± ·33	10 8·71 ± ·35	10 3·78 ± ·18	10 54·95 ± 1·54	10 56·74 ± 1·49	10 60·68 ± 2·44	10 44·70 ± 2·25
	English	50 21·47 ± ·21	50 34·78 ± ·26	50 12·09 ± ·11	50 9·38 ± ·17	50 4·68 ± ·10	50 56·78 ± ·71	50 57·97 ± ·70	50 62·72 ± 1·16	50 50·83 ± 1·17
	Egyptians	50 21·52 ± ·20	50 35·34 ± ·23	50 12·34 ± ·12	50 10·42 ± ·17	50 4·63 ± ·10	50 57·82 ± ·72	50 59·35 ± ·75	50 65·05 ± 1·25	50 44·38 ± ·69
	Veddahs	15 20·14 ± ·44	15 33·87 ± ·69	15 11·43 ± ·28	16 9·14 ± ·28	16 3·93 ± ·15	15 57·08 ± 1·25	15 61·66 ± 1·44	15 68·92 ± 2·41	16 43·58 ± 1·65
	Chimpanzees	33 18·86 ± ·33	33 22·99 ± ·42	33 5·44 ± ·15	13 8·32 ± ·34	13 1·78 ± ·14	33 28·91 ± ·66	33 30·55 ± ·67	33 22·02 ± ·84	13 21·42 ± 1·43
	Gorillas	27 22·76 ± ·52	27 30·78 ± ·55	27 9·04 ± ·21	21 6·47 ± ·30	21 3·31 ± ·11	27 40·82 ± 1·26	27 41·03 ± 1·04	27 36·43 ± 1·50	21 56·80 ± 3·26
	Orang-utans	44 15·43 ± ·24	44 33·76 ± ·57	44 11·16 ± ·20	—	—	44 74·45 ± 1·89	44 91·78 ± 2·39	44 122·78 ± 4·38	—
Both Sexes	*Gibbons* { *Hylobates*	25 10·54 ± ·15	25 16·70 ± ·29	25 3·85 ± ·13	16 5·33 ± ·13	16 0·46 ± ·04	25 36·88 ± 1·18	25 55·48 ± 1·36	25 58·87 ± 2·22	16 8·66 ± ·75
	Gibbons { *Symphalangus*	9 13·17	9 18·53	9 3·89	6 6·00	6 0·62	9 28·98	9 43·20	9 39·77	6 10·15
Females	Congo	27 23·27 ± ·26	27 30·81 ± ·32	27 8·63 ± ·16	25 9·72 ± ·35	25 2·41 ± ·12	27 37·22 ± ·67	27 38·59 ± ·60	27 32·60 ± ·82	25 25·70 ± 1·37
	Gaboon, 1864	37 22·52 ± ·32	37 31·81 ± ·37	37 9·82 ± ·17	35 9·88 ± ·29	35 2·68 ± ·11	37 44·01 ± ·80	37 44·87 ± ·84	37 42·04 ± 1·27	35 27·71 ± ·97
	Gaboon, 1880	18 21·84 ± ·33	18 31·00 ± ·50	18 9·60 ± ·24	18 9·46 ± ·27	18 2·13 ± ·13	18 44·14 ± 1·11	18 45·09 ± 1·21	18 42·34 ± 1·88	18 22·86 ± 1·49
	Philippine Islands ...	5 19·36	5 28·00	5 8·66	4 8·80	4 2·58	5 44·84	5 47·00	5 45·04	4 29·00
	Borneo	12 19·68 ± ·33	12 27·88 ± ·43	12 8·00 ± ·20	11 7·66 ± ·40	11 2·59 ± ·19	12 41·07 ± 1·35	12 44·93 ± 1·06	12 41·92 ± 1·61	11 34·71 ± 2·06
	Javanese	6 19·42	6 27·58	6 8·18	6 8·58	6 2·80	6 42·75	6 45·55	6 43·10	6 32·67
	Nubians*	19 23·37	19 32·08	19 9·22	19 11·90	19 2·81	19 40·32	19 42·82	19 38·87	19 23·48
	Hindoos	5 18·90	5 29·00	5 9·24	5 7·86	5 3·02	5 49·38	5 52·50	5 53·74	5 39·06
	English	50 20·55 ± ·19	50 31·50 ± ·28	50 10·81 ± ·12	50 9·10 ± ·17	50 4·16 ± ·07	50 52·92 ± ·63	50 52·45 ± ·61	50 53·70 ± ·98	50 46·58 ± ·77
	Egyptians	50 20·63 ± ·25	50 32·04 ± ·33	50 10·78 ± ·12	50 10·84 ± ·16	50 3·97 ± ·09	50 52·74 ± ·63	50 53·86 ± ·69	50 56·02 ± 1·12	50 36·94 ± ·79
	Veddahs	12 18·52 ± ·44	12 31·83 ± ·59	12 10·70 ± ·29	12 8·28 ± ·33	12 3·08 ± ·20	12 59·24 ± 2·44	12 64·23 ± 2·11	12 73·38 ± 3·65	12 36·86 ± 2·62
	Chimpanzees	18 17·63 ± ·54	18 21·45 ± ·63	18 5·09 ± ·20	9 8·07 ± ·18	9 1·47 ± ·13	18 29·09 ± ·96	18 30·50 ± 1·24	18 22·29 ± 1·54	9 18·53 ± 1·73
	Gorillas	25 20·22 ± ·51	25 28·58 ± ·62	25 8·88 ± ·21	19 5·31 ± ·23	19 2·82 ± ·16	25 44·87 ± 1·16	25 44·94 ± 1·28	25 42·74 ± 1·93	19 55·07 ± 3·15
	Orang-utans	13 11·85 ± ·26	13 27·75 ± ·65	13 9·29 ± ·22	—	—	13 79·24 ± 2·48	13 99·47 ± 3·51	13 136·55 ± 6·50	—

* Mean of results obtained by two sexings.

the Chimpanzee, although all the races dealt with are nearer to the Gorilla than either of the two apes. The following table gives the racial order for the simotic and two mesodacryal indices:

TABLE IX.

Simotic and Mesodacryal Indices. Mean Values.*

Simotic Index (S)				Mesodacryal (Subtense) Index (α)				Mesodacryal (Arcual) Index (β)			
♂		♀		♂		♀		♂		♀	
[*Orang-utans*	0?]	[*Orang-utans*	17?]	—		—		—		—	
Chimpanzees	21	*Chimpanzees*	19	*Chimpanzees*	29	*Chimpanzees*	29	*Chimpanzees*	22	*Chimpanzees*	22
Congo	26	Nubians†	23	Congo	39	Congo	37	Congo	35	Congo	33
Moluccas	27	—		*Gorillas*	41	Nubians†	40	*Gorillas*	36	—	
Gaboon, 1880	29	Gaboon, 1880	23	Philippines	42	Borneo	41	Philippines	38	—	
Sumatra	29	Congo	26	Borneo	42	Javanese	43	Gaboon, 1864	40	Nubians†	39
Celebes	30	—		Moluccas	43	Gaboon, 1880	44	Moluccas	41	Gaboon, 1864	42
Borneo	30	—		Gaboon, 1864	43	Gaboon, 1864	44	Gaboon, 1880	42	Gaboon, 1880	42
Philippines	30	Gaboon, 1864	28	Gaboon, 1880	43	*Gorillas*	45	Borneo	42	Borneo	42
Gaboon, 1864	31	Philippines	29	Malays	44	Philippines	45	Malays	43	*Gorillas*	43
Malays	32	—		Celebes	45	—		Celebes	45	Javanese	43
Javanese	34	Javanese	33	Sumatra	45	—		Javanese	46	Philippines	45
Nubians†	34	Borneo	35	Javanese	45	—		Sumatra	47	—	
Ainos	43	—		Nubians†	48	—		Ainos	50	—	
Veddahs	44	Veddahs	37	Ainos	50	—		Nubians†	51	—	
Egyptians	44	Egyptians	37	Hindoos	55	Hindoos	49	Hindoos	61	Hindoos	54
Hindoos	45	Hindoos	39	English	57	English	53	English	63	English	54
English	51	English	47	Veddahs	57	Egyptians	53	Egyptians	65	Egyptians	56
Gorillas	57	*Gorillas*	55	Egyptians	58	Veddahs	59	Veddahs	69	Veddahs	73
—		—		*Orang-utans*	74	*Orang-utans*	79	*Orang-utans*	123	*Orang-utans*	137

Now if we compare the results given in this table we see that with one exception, that of the Borneo crania, the simotic index is greater in the males than in the females; but there is a greater agreement between the sexes—the Hindoos, English and Egyptians excepted—in the mesodacryal indices. For the mesodacryal index, α, there are 7 cases (judging from Table VIII, p. 403) in which the female is greater, as against 7 in which the male is greater; for the mesodacryal index β, there are also 7 cases for the male and 7 for the female with greater value, and

* The following are the values for the Gibbons:

		Simotic Index	Mesodacryal (Subtense) Index	Mesodacryal (Arcual) Index
Gibbons	*Hylobates* ♂ + ♀ + δ	9	37	55
	Symphalangus ♂ + ♀ + δ	10	29	43

† The position of the "Nubians" confirms the view that the men were more Egyptian and the women almost negroes.

the individual cases are the same for α and β. On the whole it is the lower races for which the female has the greater value, and the higher races—in the scale of civilisation—for which the male has the greater value. Thus the mesodacryal index appears more of a racial and less of a sexual character than the simotic index. Roughly also the two sexes give nearly the same racial order, especially if we regard the large probable errors (see Table VIII, p. 403) of some of our results. Looked at in broad outline we have the nose of the Gorilla closely associated with that of the Negro, and the Negro lies between the Chimpanzee and the Asiatic races. At the bottom of the scale we have the Orang-utans closest of all to the Veddahs, who lie between the Orang-utans and the higher races, so far as such are provided by these measurements*. On the whole the racial scale of the mesodacryal indices seems far more suggestive than that of the simotic index, where the Gorilla is widely separated from the Negro, and is divided from the Veddahs by English, Hindoos and Egyptians†. We believe that a measure of the mesodacryal index is likely to be of as much value as the determination of the simotic index, possibly of greater value.

(5) *On the Absolute Nasal Measurements.*

We are doubtful as to what if any stress can be laid on the absolute measurements. Mérejkowsky, however, cites the absolute values of the simotic subtenses and it is of interest to compare our results with his. This is done in Table X. It will be seen that for Negroes, Malays and "White races" our results are in fair agreement with Mérejkowsky's; with his Melanesian, Polynesian and American groups, we have nothing to compare. There is relatively little difference in the order between the two sexes, but the actual height of the nasal bones, with the one exception of the Borneo crania, is considerably less for the females. This smaller value of the simotic subtense in the female is not true for the chord, which in Nubians, Egyptians, Gaboon (1864), Gaboon (1880), and Philippine Islanders is greater in the female as will be seen in Table XI. There is very little we think to be learnt concerning racial differences in either Table X or Table XI.

The races are mixed in a rather meaningless way—at least we have failed to find any clue to their order, although in this respect the simotic chord appears worse than the simotic subtense. On the whole we believe little of real value is to be learnt from the absolute size of the nasal bones.

* The Gibbons agree with the Orang for the simotic index, but are nearer to the Chimpanzee than to the Orang for the mesodacryal index α, while for the arcual index β their position, like that of the Gorilla, is within the human range.

† Judged by the mesodacryal indices the Gorilla has a flatter nose than any human race, the Negroes approaching most closely; judged by the simotic index the Gorilla has a sharper nose than any human race, the English approaching most closely. At any rate from the standpoint of physiognomy, there can be no doubt that the mesodacryal indices express better the general visual state of affairs: see, however, the contours, pp. 428—9.

TABLE X. *Simotic Subtense* *.

Males		Females		Sex ?† Mérejkowsky	
Chimpanzees	1·78	*Chimpanzees*	1·47	—	
Gaboon, 1880	2·30	Gaboon, 1880	2·13	—	
Borneo	2·33	Congo	2·41	—	
Congo	2·45	Philippines	2·58	Negroes	2·60
Moluccas	2·48	—		—	
Philippines	2·55	Borneo	2·59	Mongols	3·00
Celebes	2·58	—		—	
Malays	2·62	—		Malays	2·90
Gaboon, 1864	2·88	Gaboon, 1864	2·68	—	
Sumatra	2·88	—		Polynesians	3·13
Javanese	2·99	Javanese	2·80	—	
Gorillas	3·31	Nubians	2·81	—	
Nubians	3·74	*Gorillas*	2·82	—	
Hindoos	3·78	Hindoos	3·02	—	
Ainos	3·80	—		American Indians	3·80
Veddahs	3·93	Veddahs	3·08	Melanesians	3·85
Egyptians	4·63	Egyptians	3·97	—	
English	4·68	English	4·16	White races	4·74

TABLE XI. *Simotic Chord* ‡.

Males		Females	
Gorillas	6·47	*Gorillas*	5·31
Malays	7·76	—	
Borneo	7·93	Borneo	7·66
Gaboon, 1880	8·30	Hindoos	7·86
Chimpanzees	8·32	*Chimpanzees*	8·07
Philippines	8·42	Veddahs	8·28
Celebes	8·54	—	
Hindoos	8·71	Javanese	8·58
Ainos	8·80	—	
Javanese	8·83	Philippines	8·80
Veddahs	9·14	—	
Moluccas	9·36	English	9·10
English	9·38	—	
Gaboon, 1864	9·43	Gaboon, 1880	9·46
Congo	9·80	Congo	9·72
Sumatra	10·10	Gaboon, 1864	9·88
Egyptians	10·42	Egyptians	10·84
Nubians	11·45	Nubians	11·90

* The Gibbons lie at the very top of this scale close to the Orang-utan with its practically zero value of the subtense. Thus for the simotic subtense we have:

$$\textit{Gibbons},\ ♂ + ♀ + ⚲ \begin{cases} \textit{Hylobates} & 0{\cdot}46 \\ \textit{Symphalangus} & 0{\cdot}62 \end{cases}$$

† Apparently Mérejkowsky mixes both sexes together, which is hardly permissible considering the large sexual differences.

‡ $$\textit{Gibbons},\ ♂ + ♀ + ⚲ \begin{cases} \textit{Hylobates} & 5{\cdot}33 \\ \textit{Symphalangus} & 6{\cdot}00 \end{cases}$$

(6) *On the Absolute Mesodacryal Measurements.*

The following table gives the order for the mesodacryal subtense, chord and arc:

TABLE XII.

Racial Order of Mesodacryal Measurements.

Subtense				Chord				Arc			
♂		♀		♂		♀		♂		♀	
Chimpanzees	5·44	*Chimpanzees*	5·09	*Orangs*	15·4	*Orangs*	11·9	*Chimpanzees*	23·0	*Chimpanzees*	21·4
Philippines	8·42	Borneo	8·00	*Chimpanzees*	18·9	*Chimpanzees*	17·6	Philippines	28·1	Javanese	27·6
Borneo	8·86	Javanese	8·18	Hindoos	19·5	Veddahs	18·5	Malays	30·0	—	
Gorillas	9·04	Congo	8·63	Veddahs	20·1	Hindoos	18·9	Borneo	30·1	*Orangs*	27·8
Moluccas	9·15	—		Philippines	20·3	Philippines	19·4	Moluccas	30·2	Borneo	27·9
Malays	9·23	—		Ainos	20·9	—		Celebes	30·8	—	
Congo	9·32	Philippines	8·66	Malays	21·1	—		*Gorillas*	30·8	Philippines	28·0
Celebes	9·40	—		Borneo	21·2	Javanese	19·4	Hindoos	31·2	*Gorillas*	28·6
Javanese	9·63	*Gorillas*	8·88	Celebes	21·3	—		Ainos	31·4	—	
Gaboon (1880)	9·78	Nubians	9·22	Moluccas	21·4	—		Javanese	31·4	Hindoos	29·0
Sumatra	9·86	—		English	21·5	Borneo	19·7	Gaboon (1864)	32·4	Congo	30·8
Gaboon (1864)	9·91	Hindoos	9·24	Egyptians	21·5	*Gorillas*	20·2	Congo	32·4	Gaboon (1880)	31·0
Ainos	10·42	—		Javanese	21·6	English	20·6	Sumatra	32·6	—	
Hindoos	10·66	*Orangs*	9·29	Sumatra	22·0	—		Gaboon (1880)	32·6	—	
Orangs	11·16	Gaboon (1880)	9·60	*Gorillas*	22·8	Egyptians	20·6	*Orangs*	33·8	English	31·5
Veddahs	11·43	Gaboon (1864)	9·82	Gaboon (1864)	23·2	Gaboon (1880)	21·8	Veddahs	33·9	Veddahs	31·8
Nubians	11·72	Veddahs	10·70	Gaboon (1880)	23·2	Gaboon (1864)	22·5	English	34·8	Gaboon (1864)	31·8
English	12·09	Egyptians	10·78	Congo	23·9	Congo	23·3	Egyptians	35·3	Egyptians	32·0
Egyptians	12·34	English	10·81	Nubians	24·2	Nubians	23·4	Nubians	36·5	Nubians	32·1
Gibbons, ♂ + ♀ + ☿		{ *Hylobates*	3·85	*Gibbons*, ♂ + ♀ + ☿		{ *Hylobates*	10·54	*Gibbons*, ♂ + ♀ + ☿		{ *Hylobates*	16·70
		{ *Symphalangus*	3·89			{ *Symphalangus*	13·12			{ *Symphalangus*	18·53

Now we have already seen that the sexual differences in the mesodacryal indices are not very large, but we notice at once that these sexual differences are very considerable in the absolute measurements. It becomes rather meaningless to assert that the Hindoo's eyes are closer together than the Englishman's, when the difference in mesodacryal chord is less than that between male and female Javanese. They are less for the same reason, namely, because the cranium as a whole is smaller. Practically the female is less—and often very considerably less—in every mesodacryal absolute measurement, and we cannot usefully compare males and females—except as to absolute size, a sexual difference we are already familiar with—unless we get rid of this problem of size by using as we have already done the indices. But surely if this be true for sex, it is also true for race. Our scales become scales of absolute size and of very small racial significance. In all probability absolute lengths or breadths of the head would tell just as much as is to be learnt from these tables of relative racial sizes of the bridge of the nose. In our opinion there is little that has bearing on racial relationships. In, for example, the chord scales, it is possible that Veddah and Orang-utans are close together because of an ancestral link, but the Hindoo appears in this part of the scale because of his small size. The association of the Negroes with the Gorillas may again be partly ancestral, but the association of English and Javanese in the same part of the scale as the Gorillas is probably fortuitous. We are inclined to say that very little indeed can be deduced from absolute scales of this kind for nasal characters.

It is often asserted that a characteristic feature of the negro lies in the fact that he has his eyes farther apart than other races of man. The average value of the mesodacryal chord for the Congo and Gaboon male negroes is 23·5 and for English and Egyptians 21·5; for the females, Congo and Gaboon, it is 22·6, and for English and Egyptians 20·6. There is thus 2 mm. difference on the average. Is it possible for the human eye to appreciate this difference? We are inclined to doubt it and believe that it is the marked simotic platygephyrosis (see p. 430 below) of the negro nose which produces largely the impression of greater ocular

Value of Ratio 100 × *Mesodacryal Chord/Minimum Forehead Breadth.*

Race	♂	♀
Congo	24·6	25·1
Gaboon, 1864 ...	24·0	24·3
Gaboon, 1880 ...	23·5	23·5
Egyptians ...	22·6	22·4
Veddahs	22·0	21·2
English	21·9	22·1
Hindoos	21·3	21·0

breadth*. As the total size of the skull varies much from race to race it seems desirable to compare the mesodacryal chord with another horizontal measure and we have taken the minimum forehead breadth: see Table, p. 408.

It will be seen that the Negroes stand at one end of the list and the Indo-Europeans at the other, but again the change of index is relatively small, and it may be legitimately doubted whether it would be visually significant.

(7) *On the relative Racial Variability of the Nasal Bridge Characters.*

Tables XIII and XIV give respectively the Standard Deviations and the Coefficients of Variation of the characters determined in such series as supplied sufficient data for even an attempt at these constants. There is no doubt in our minds that the coefficient of variation is the better measure for racial purposes in the case of absolute measurements. We will consider first what points arise from a discussion of the variability of the mesodacryal subtense, chord and arc from this standpoint. The relative orders of variability are given in Table XV.

Again beyond the broad fact that the apes on the whole are far more variable than the races of men, perhaps little can be learnt from this table. In the apes the female is more variable than the male in 5 out of 9 cases; the same ratio of 5 to 4 is maintained in the case of the negro races; in the lower races from Borneo and the Veddahs, the males are more variable than the females in the ratio of 5 cases to 1, while in the higher races—Egyptian and English—the ratio is 5 cases of the female to 1 case of the male more variable. Altogether therefore the female is more variable in 16 and the male in 14 cases. Generally there is no evidence for greater male variation in these coefficients of variation of the mesodacryal absolute lengths. Even if we based our comparison on the absolute variations of the mesodacryal characters, we find among the apes the male is more variable in 5 the female in 3, with one case—the Chimpanzee mesodacryal subtense—of equality; in the lower races (Veddahs and Borneo) the male is more variable in all 6 cases; in the negro races in 5 out of 9, but in the higher races (English and Egyptian) in only 1 out of 6 cases. Thus in absolute measurement the male is more variable in 17, the female in 12 cases.

* The reader must not judge from Figs. 6—9 of the contours, pp. 428—9. These all have the mesodacryal chord in considerable excess of the negro mean, and were selected to give approximately mean *angles*, not mean absolute lengths.

TABLE XIII. *Man and Anthropoid Apes. Nasal Bridge Characters. Standard Deviations.*

		Mesodacryal Chord	Mesodacryal Arc	Mesodacryal Subtense	Simotic Chord	Simotic Subtense	Index a	Index a'	Index β	Index S
Males	Congo	*50* 2·70 ± ·18	*50* 4·02 ± ·27	*50* 1·53 ± ·10	*50* 2·24 ± ·15	*50* ·91 ± ·06	*50* 5·67 ± ·38	*50* 5·92 ± ·40	*50* 8·67 ± ·58	*50* 7·64 ± ·52
	Gaboon, 1864 ...	*48* 1·89 ± ·13	*48* 2·86 ± ·20	*48* 1·37 ± ·09	*47* 2·66 ± ·18	*47* 1·08 ± ·07	*48* 5·65 ± ·39	*48* 5·22 ± ·36	*48* 7·75 ± ·53	*47* 8·82 ± ·61
	Gaboon, 1880 ...	*16* 2·70 ± ·32	*16* 2·37 ± ·28	*16* 1·11 ± ·13	*16* 2·38 ± ·28	*16* ·67 ± ·08	*16* 7·63 ± ·91	*16* 7·43 ± ·89	*16* 11·32 ± 1·35	*16* 8·81 ± 1·05
	Philippine Islands	*17* 1·76 ± ·20	*17* 2·15 ± ·25	*17* 1·47 ± ·17	*17* 1·91 ± ·22	*17* ·80 ± ·09	*17* 6·60 ± ·76	*17* 5·08 ± ·59	*17* 7·68 ± ·89	*17* 6·64 ± ·77
	Borneo	*26* 2·39 ± ·22	*26* 2·43 ± ·23	*26* 1·08 ± ·10	*26* 1·81 ± ·17	*26* ·77 ± ·07	*26* 8·01 ± ·75	*26* 7·85 ± ·73	*26* 12·49 ± 1·17	*26* 9·35 ± ·88
	Malays	*14* 2·67 ± ·34	*14* 3·53 ± ·45	*14* 1·24 ± ·16	*14* 1·93 ± ·25	*14* ·95 ± ·12	*14* 6·34 ± ·81	*14* 6·90 ± ·88	*14* 10·55 ± 1·34	*14* 7·81 ± 1·00
	Java and Madura	*39* 2·10 ± ·16	*39* 3·05 ± ·23	*39* 1·68 ± ·13	*39* 2·04 ± ·16	*39* 1·15 ± ·09	*39* 8·74 ± ·67	*39* 8·81 ± ·67	*39* 13·72 ± 1·05	*39* 11·20 ± ·86
	Hindoos	*10* 2·04 ± ·31	*10* 3·20 ± ·48	*10* 1·54 ± ·23	*10* 1·62 ± ·24	*10* ·83 ± ·13	*10* 7·23 ± 1·09	*10* 6·97 ± 1·05	*10* 11·46 ± 1·73	*10* 10·55 ± 1·59
	English	*50* 2·16 ± ·15	*50* 2·71 ± ·18	*50* 1·20 ± ·08	*50* 1·77 ± ·12	*50* 1·04 ± ·07	*50* 7·47 ± ·50	*50* 7·32 ± ·49	*50* 12·19 ± ·82	*50* 12·21 ± ·82
	Egyptians ...	*50* 2·05 ± ·14	*50* 2·45 ± ·17	*50* 1·21 ± ·08	*50* 1·75 ± ·12	*50* 1·07 ± ·07	*50* 7·57 ± ·51	*50* 7·89 ± ·53	*50* 13·07 ± ·88	*50* 7·27 ± ·49
	Veddahs ...	*15* 2·51 ± ·31	*15* 3·95 ± ·49	*15* 1·61 ± ·20	*16* 1·68 ± ·20	*16* ·90 ± ·11	*15* 7·16 ± ·88	*15* 8·24 ± 1·01	*15* 13·86 ± 1·71	*16* 9·77 ± 1·16
	Chimpanzees ...	*33* 2·81 ± ·23	*33* 3·61 ± ·30	*33* 1·27 ± ·11	*13* 1·82 ± ·24	*13* ·73 ± ·10	*33* 5·65 ± ·47	*33* 5·74 ± ·48	*33* 7·13 ± ·59	*13* 7·64 ± 1·01
	Gorillas	*27* 3·97 ± ·36	*27* 4·23 ± ·39	*27* 1·66 ± ·15	*21* 2·01 ± ·21	*21* ·73 ± ·08	*27* 9·69 ± ·89	*27* 8·00 ± ·73	*27* 11·56 ± 1·06	*21* 22·15 ± 2·30
	Orang-utans ...	*44* 2·35 ± ·17	*44* 5·56 ± ·40	*44* 1·93 ± ·14	—	—	*44* 18·63 ± 1·34	*44* 23·49 ± 1·69	*44* 43·04 ± 3·09	—
Both Sexes	*Gibbons, Hylobates*	*25* 1·13 ± ·11	*25* 2·12 ± ·20	*25* 0·85 ± ·08	*16* 0·77 ± ·09	*16* 0·26 ± ·03	*25* 8·74 ± ·83	*25* 10·06 ± ·96	*25* 16·43 ± 1·57	*16* 4·48 ± ·53
Females	Congo	*27* 2·00 ± ·18	*27* 2·43 ± ·22	*27* 1·20 ± ·11	*25* 2·58 ± ·25	*25* ·90 ± ·09	*27* 5·15 ± ·47	*27* 4·58 ± ·42	*27* 6·35 ± ·58	*25* 10·15 ± ·97
	Gaboon, 1864 ...	*37* 2·86 ± ·22	*37* 3·35 ± ·26	*37* 1·51 ± ·12	*35* 2·57 ± ·21	*35* ·92 ± ·07	*37* 7·19 ± ·56	*37* 7·61 ± ·60	*37* 11·49 ± ·90	*35* 8·47 ± ·68
	Gaboon, 1880 ...	*18* 2·10 ± ·24	*18* 3·15 ± ·35	*18* 1·52 ± ·17	*18* 1·69 ± ·19	*18* ·83 ± ·09	*18* 7·00 ± ·78	*18* 7·60 ± ·85	*18* 11·84 ± 1·33	*18* 9·38 ± 1·05
	Borneo	*12* 1·69 ± ·23	*12* 2·23 ± ·31	*12* 1·01 ± ·14	*11* 1·94 ± ·28	*11* ·95 ± ·14	*12* 6·94 ± ·96	*12* 5·46 ± ·75	*12* 8·24 ± 1·13	*11* 10·15 ± 1·46
	English	*50* 2·04 ± ·14	*50* 2·90 ± ·20	*50* 1·27 ± ·09	*50* 1·81 ± ·12	*50* ·78 ± ·05	*50* 6·64 ± ·45	*50* 6·36 ± ·43	*50* 10·25 ± ·69	*50* 8·09 ± ·55
	Egyptians ...	*50* 2·59 ± ·17	*50* 3·47 ± ·23	*50* 1·22 ± ·08	*50* 1·66 ± ·11	*50* ·90 ± ·06	*50* 6·62 ± ·45	*50* 7·19 ± ·49	*50* 11·76 ± ·79	*50* 8·24 ± ·56
	Veddahs	*12* 2·27 ± ·31	*12* 3·04 ± ·42	*12* 1·49 ± ·21	*12* 1·72 ± ·24	*12* 1·03 ± ·14	*12* 12·52 ± 1·72	*12* 10·86 ± 1·49	*12* 18·77 ± 2·58	*12* 13·45 ± 1·85
	Chimpanzees ...	*18* 3·38 ± ·38	*18* 3·93 ± ·44	*18* 1·27 ± ·14	*9* ·82 ± ·13	*9* ·58 ± ·09	*18* 6·05 ± ·68	*18* 7·77 ± ·87	*18* 9·70 ± 1·09	*9* 7·69 ± 1·22
	Gorillas	*25* 3·75 ± ·36	*25* 4·57 ± ·44	*25* 1·59 ± ·15	*19* 1·51 ± ·17	*19* 1·04 ± ·11	*25* 8·61 ± ·82	*25* 9·47 ± ·90	*25* 14·27 ± 1·36	*19* 20·33 ± 2·22
	Orang-utans ...	*13* 1·37 ± ·18	*13* 3·49 ± ·46	*13* 1·20 ± ·16	—	—	*13* 13·27 ± 1·76	*13* 18·78 ± 2·48	*13* 34·72 ± 4·59	—

TABLE XIV. *Man and Anthropoid Apes. Nasal Bridge Characters. Coefficients of Variation.*

		Mesodacryal Chord	Mesodacryal Arc	Mesodacryal Subtense	Simotic Chord	Simotic Subtense	Index α	Index α'	Index β	Index S
Males	Congo	*50* 11·28 ± ·77	*50* 12·41 ± ·85	*50* 16·44 ± 1·14	*50* 22·89 ± 1·62	*50* 36·85 ± 2·80	*50* 14·49 ± 1·00	*50* 14·61 ± 1·01	*50* 24·43 ± 1·74	*50* 29·87 ± 2·19
	Gaboon, 1864 ...	*48* 8·15 ± ·56	*48* 8·83 ± ·61	*48* 13·84 ± ·97	*47* 28·15 ± 2·11	*47* 37·35 ± 2·94	*48* 13·16 ± ·92	*48* 11·98 ± ·84	*48* 19·39 ± 1·38	*47* 28·54 ± 2·14
	Gaboon, 1880 ...	*16* 11·67 ± 1·41	*16* 7·28 ± ·87	*16* 11·33 ± 1·37	*16* 28·66 ± 3·69	*16* 29·29 ± 3·78	*16* 17·77 ± 2·18	*16* 16·57 ± 2·03	*16* 26·99 ± 3·44	*16* 30·37 ± 3·94
	Philippine Islands	*17* 8·67 ± 1·01	*17* 7·63 ± ·89	*17* 17·42 ± 2·08	*17* 22·68 ± 2·76	*17* 31·27 ± 3·95	*17* 15·89 ± 1·88	*17* 11·92 ± 1·40	*17* 19·97 ± 2·40	*17* 21·84 ± 2·64
	Borneo	*26* 11·23 ± 1·06	*26* 8·09 ± ·76	*26* 12·22 ± 1·16	*26* 22·78 ± 2·24	*26* 32·86 ± 3·39	*26* 18·97 ± 1·84	*26* 17·45 ± 1·68	*26* 29·61 ± 3·00	*26* 31·03 ± 3·17
	Malays	*14* 12·64 ± 1·64	*14* 11·77 ± 1·52	*14* 13·45 ± 1·75	*14* 24·81 ± 3·35	*14* 36·32 ± 5·20	*14* 14·36 ± 1·87	*14* 15·24 ± 1·99	*14* 24·82 ± 3·35	*14* 23·16 ± 3·11
	Java and Madura	*39* 9·72 ± ·75	*39* 9·73 ± ·75	*39* 17·46 ± 1·37	*39* 23·06 ± 1·85	*39* 38·38 ± 3·34	*39* 19·45 ± 1·54	*39* 18·70 ± 1·48	*39* 30·10 ± 2·50	*39* 33·06 ± 2·79
	Hindoos	*10* 10·49 ± 1·60	*10* 10·27 ± 1·56	*10* 14·48 ± 2·23	*10* 18·60 ± 2·90	*10* 21·94 ± 3·46	*10* 13·16 ± 2·02	*10* 12·29 ± 1·88	*10* 18·88 ± 2·95	*10* 23·60 ± 3·75
	English	*50* 10·04 ± ·68	*50* 7·80 ± ·53	*50* 9·93 ± ·68	*50* 18·85 ± 1·32	*50* 22·13 ± 1·56	*50* 13·16 ± ·90	*50* 12·63 ± ·87	*50* 19·43 ± 1·36	*50* 24·03 ± 1·71
	Egyptians ...	*50* 9·51 ± ·65	*50* 6·93 ± ·47	*50* 9·80 ± ·67	*50* 16·84 ± 1·17	*50* 23·10 ± 1·64	*50* 13·09 ± ·90	*50* 13·29 ± ·91	*50* 20·09 ± 1·41	*50* 16·38 ± 1·13
	Veddahs	*15* 12·45 ± 1·56	*15* 11·67 ± 1·46	*15* 14·12 ± 1·77	*16* 18·35 ± 2·26	*16* 23·04 ± 2·89	*15* 12·54 ± 1·57	*15* 13·37 ± 1·68	*15* 20·11 ± 2·57	*16* 22·42 ± 2·80
	Chimpanzees ...	*33* 14·91 ± 1·27	*33* 15·72 ± 1·34	*33* 23·34 ± 2·04	*13* 21·88 ± 3·03	*13* 41·07 ± 6·28	*33* 19·53 ± 1·68	*33* 18·80 ± 1·62	*33* 32·39 ± 2·96	*13* 35·69 ± 5·29
	Gorillas	*27* 17·46 ± 1·65	*27* 13·75 ± 1·29	*27* 18·31 ± 1·74	*21* 31·12 ± 3·54	*21* 21·96 ± 2·39	*27* 23·74 ± 2·30	*27* 19·51 ± 1·86	*27* 31·74 ± 3·19	*21* 38·99 ± 4·63
	Orang-utans ...	*44* 15·22 ± 1·12	*44* 16·46 ± 1·22	*44* 17·28 ± 1·28	—	—	*44* 25·02 ± 1·91	*44* 25·60 ± 1·96	*44* 35·06 ± 2·81	—
Both Sexes	*Gibbons, Hylobates*	*25* 10·67 ± 1·03	*25* 12·69 ± 1·23	*25* 22·05 ± 2·20	*16* 14·49 ± 1·76	*16* 57·15 ± 8·76	*25* 23·70 ± 2·38	*25* 18·13 ± 1·79	*25* 27·90 ± 2·86	*16* 51·72 ± 7·64
Females	Congo	*27* 8·61 ± ·80	*27* 7·89 ± ·73	*27* 13·86 ± 1·30	*25* 26·58 ± 2·71	*25* 37·33 ± 4·03	*27* 13·84 ± 1·29	*27* 11·88 ± 1·11	*27* 19·48 ± 1·85	*25* 39·47 ± 4·31
	Gaboon, 1864 ...	*37* 12·68 ± 1·01	*37* 10·54 ± ·84	*37* 15·34 ± 1·23	*35* 25·97 ± 2·23	*35* 34·50 ± 3·09	*37* 16·34 ± 1·32	*37* 16·95 ± 1·37	*37* 27·34 ± 2·30	*35* 30·58 ± 2·69
	Gaboon, 1880 ...	*18* 9·60 ± 1·09	*18* 10·17 ± 1·16	*18* 15·78 ± 1·82	*18* 17·88 ± 2·07	*18* 39·15 ± 5·03	*18* 15·87 ± 1·83	*18* 16·85 ± 1·95	*18* 27·97 ± 3·38	*18* 41·04 ± 5·33
	Borneo	*12* 8·60 ± 1·19	*12* 7·99 ± 1·11	*12* 12·56 ± 1·76	*11* 25·35 ± 3·87	*11* 36·59 ± 5·93	*12* 16·89 ± 2·39	*12* 12·15 ± 1·70	*12* 19·66 ± 2·81	*11* 29·24 ± 4·55
	English	*50* 9·93 ± ·68	*50* 9·21 ± ·63	*50* 11·74 ± ·80	*50* 19·85 ± 1·39	*50* 18·83 ± 1·31	*50* 12·55 ± ·86	*50* 12·14 ± ·83	*50* 19·09 ± 1·33	*50* 17·36 ± 1·21
	Egyptians ...	*50* 12·57 ± ·86	*50* 10·83 ± ·74	*50* 11·32 ± ·77	*50* 15·28 ± 1·05	*50* 22·78 ± 1·61	*50* 12·56 ± ·86	*50* 13·35 ± ·92	*50* 21·00 ± 1·48	*50* 22·32 ± 1·58
	Veddahs	*12* 12·28 ± 1·72	*12* 9·56 ± 1·33	*12* 13·95 ± 1·96	*12* 20·71 ± 2·97	*12* 33·44 ± 5·09	*12* 21·13 ± 3·04	*12* 16·90 ± 2·39	*12* 25·57 ± 3·74	*12* 36·48 ± 5·65
	Chimpanzees ...	*18* 19·19 ± 2·24	*18* 18·33 ± 2·13	*18* 24·95 ± 2·97	*9* 10·11 ± 1·62	*9* 39·22 ± 7·13	*18* 20·79 ± 2·44	*18* 25·48 ± 3·05	*18* 43·51 ± 5·74	*9* 41·48 ± 7·65
	Gorillas	*25* 18·56 ± 1·83	*25* 16·00 ± 1·57	*25* 17·91 ± 1·76	*19* 28·40 ± 3·35	*19* 36·98 ± 4·57	*25* 19·19 ± 1·90	*25* 21·07 ± 2·10	*25* 33·40 ± 3·52	*19* 36·91 ± 4·56
	Orang-utans ...	*13* 11·54 ± 1·55	*13* 12·58 ± 1·69	*13* 12·88 ± 1·73	—	—	*13* 16·75 ± 2·28	*13* 18·88 ± 2·59	*13* 25·43 ± 3·57	—

TABLE XV.

*Variability Scales for Mesodacryal Characters**.

Subtense				Chord				Arc			
♂		♀		♂		♀		♂		♀	
Egyptians	9·8	Egyptians	11·3	Gaboon (1864)	8·2	Borneo	8·6	Egyptians	6·9	Congo	7·9
English	9·9	English	11·7	Philippines	8·7	—		Gaboon (1880)	7·3	Borneo	8·0
Gaboon (1880)	11·3	Borneo	12·6	Egyptians	9·5	Congo	8·6	Philippines	7·6	—	
Borneo	12·2	*Orangs*	12·9	Javanese	9·7	—		English	7·8	English	9·2
Malays	13·5	—		English	10·0	Gaboon (1880)	9·6	Borneo	8·1	Veddahs	9·6
Gaboon (1864)	13·8	Congo	13·9	Hindoos	10·5	—		Gaboon (1864)	8·8	Gaboon (1880)	10·2
Veddahs	14·1	Veddahs	14·0	Borneo	11·2	English	9·9	Javanese	9·7	—	
Hindoos	14·5	—		Congo	11·3	*Orangs*	11·5	Hindoos	10·3	—	
Congo	16·4	Gaboon (1864)	15·3	Gaboon (1880)	11·7	Veddahs	12·3	Veddahs	11·7	Gaboon (1864)	10·5
Philippines	17·4	—		Veddahs	12·5	Egyptians	12·6	Malays	11·8	—	
Orangs	17·3	—		Malays	12·6	Gaboon (1864)	12·7	Congo	12·4	Egyptians	10·8
Javanese	17·5	Gaboon (1880)	15·8	*Chimpanzees*	14·9	—		*Gorillas*	13·8	*Orangs*	12·6
Gorillas	18·3	*Gorillas*	17·9	*Orangs*	15·2	*Gorillas*	18·6	*Chimpanzees*	15·7	*Gorillas*	16·0
Chimpanzees	23·3	*Chimpanzees*	25·0	*Gorillas*	17·5	*Chimpanzees*	19·2	*Orangs*	16·5	*Chimpanzees*	18·3

* The values for Hylobates are for both sexes together:

Subtense 22·1 | Chord 10·7 | Arc 12·7

For the simotic chord and subtense we have:

TABLE XVI.

*Order of Racial Variability for Simotic Characters**.

Simotic Subtense				Simotic Chord			
♂		♀		♂		♀	
Hindoos	21·9	—		Egyptians	16·8	*Chimpanzees*	10·10
Gorillas	22·0	English	18·8	Veddahs	18·4	Egyptians	15·3
English	22·1	Egyptians	22·8	Hindoos	18·6	—	
Veddahs	23·0	Veddahs	33·4	English	18·9	Gaboon (1880)	17·9
Egyptians	23·1	Gaboon (1864)	34·5	*Chimpanzees*	21·9	English	19·9
Gaboon (1880)	29·3	Borneo	36·6	Philippines	22·7	—	
Philippines	31·3	—		Borneo	22·8	Veddahs	20·7
Borneo	32·9	*Gorillas*	37·0	Congo	22·9	Borneo	25·4
Malays	36·3	—		Javanese	23·1	—	
Congo	36·9	Congo	37·3	Malays	24·8	—	
Gaboon (1864)	37·4	Gaboon (1880)	39·2	Gaboon (1864)	28·2	Gaboon (1864)	26·0
Javanese	38·4	—		Gaboon (1880)	28·7	Congo	26·6
Chimpanzees	41·1	*Chimpanzees*	39·2	*Gorillas*	31·1	*Gorillas*	28·4

Again there is practically little to be learnt, the Gorillas are very variable in nasal chord and much less variable (in scale position) on the subtense, while the reverse holds for the Chimpanzees, the subtense being far more variable than the chord. The Negroes are as a rule low down on the scale, being considerably more variable than the English, Egyptians or Hindoos. In 9 cases the female and in 9 the male is more variable in simotic characters.

Lastly we pass to the indices, omitting the calculated mesodacryal subtense index α'. We find for Standard Deviations the order given in Table XVII.

Again there seems little to be learnt with regard to racial variability from these results, beyond the confirmation of what we have learnt from the earlier tables that the Orangs, and the Gorillas as a rule, have large variability; here, however,—as distinguished from the coefficients of variation of the absolute lengths—the Chimpanzee is seen to have very small variability. If, however, we take the coefficients of variation of the indices we find the Chimpanzee is associated with the other two anthropoids of our scales in high variability. Turning to the sexual differences, we find in the apes as many cases (4) of female as of male greater variability; among the human races there are 10 cases in which the female is more variable and 11 cases in which the male is more variable, or again there is little if any distinction of sex in variability.

An examination of Tables XIII and XIV will show the reader that much of the non-significance of the variation scale for racial purposes is quite possibly due

* The values for Hylobates, both sexes together, are: simotic chord 14·5 and simotic arc 57·2.

TABLE XVII.

Order of Standard Deviations of Nasal Bridge Indices.*

Mesodacryal Subtense Index				Mesodacryal Arcual Index				Simotic Index			
♂		♀		♂		♀		♂		♀	
Chimpanzees	5·65	Congo	5·15	*Chimpanzees*	7·13	Congo	6·35	Philippines	6·64	—	
Gaboon (1864)	5·65	*Chimpanzees*	6·05	Philippines	7·68	—		Egyptians	7·27	*Chimpanzees*	7·69
Congo	5·67	Egyptians	6·62	Gaboon (1864)	7·75	Borneo	8·24	Congo	7·64	English	8·09
Malays	6·34	—		Congo	8·67	*Chimpanzees*	9·70	*Chimpanzees*	7·64	Egyptians	8·24
Philippines	6·60	—		Malays	10·55	—		Malays	7·81	—	
Veddahs	7·16	English	6·64	Gaboon (1880)	11·32	English	10·25	Gaboon (1880)	8·81	Gaboon (1864)	8·47
Hindoos	7·23	—		Hindoos	11·46	—		Gaboon (1864)	8·82	Gaboon (1880)	9·38
English	7·47	Borneo	6·94	*Gorillas*	11·56	Gaboon (1864)	11·49	Borneo	9·35	Congo	10·15
Egyptians	7·57	Gaboon (1880)	7·00	English	12·19	Egyptians	11·76	Veddahs	9·77	Borneo	10·15
Gaboon (1880)	7·63	Gaboon (1864)	7·19	Borneo	12·49	Gaboon (1880)	11·84	Hindoos	10·55	—	
Borneo	8·01	*Gorillas*	8·61	Egyptians	13·07	*Gorillas*	14·27	Javanese	11·20	—	
Javanese	8·74	—		Javanese	13·72	—		English	12·21	Veddahs	13·45
Gorillas	9·69	Veddahs	12·52	Veddahs	13·86	Veddahs	18·77	*Gorillas*	22·15	*Gorillas*	20·33
Orangs	18·63	*Orangs*	13·27	*Orangs*	43·04	*Orangs*	34·72	—		—	

* The values for Hylobates, both sexes together, are: Mesodacryal (Subtense) Index 8·7, Mesodacryal (Arcual) Index 10·1, Simotic Index 51·7.

to the smallness of our samples. Except in the case of Egyptians and English, the size of these samples was not under our control; we set out to compare the African and Asiatic anthropoids with the human races of the same environment, and we measured all the crania available, and calculated the variabilities for all series of 10 and over*. Further the homogeneity of some of the series, as the Malay and Javanese, is possibly open to doubt. We have accordingly clubbed together our variation results for the coefficients of variation of the absolute measurements, with the results given in Table XVIII.

TABLE XVIII.

Average Values of Variability of Absolute Measurements on the Nasal Bridge.

With Simotic Lengths			Without Simotic Lengths		
	♂	♀		♂	♀
Chimpanzees ...	23·4	22·4	*Chimpanzees* ...	18·0	20·8
Gorillas ...	20·5	23·6	*Gorillas* ...	16·5	17·5
Orangs ...	—	—	*Orangs* ...	16·3	12·3
[*Hylobates* ♂ + ♀ 23·4]			[*Hylobates* ♂ + ♀ 15·1]		
Mean 22·5 †			*Mean 16·9* †		
Malays ...	19·8	—	Malays ...	12·6	—
Javanese ...	19·7	—	Javanese ...	12·3	—
Mean 19·8			*Mean 12·5*		
Congo	20·0	18·9	Congo	13·5	10·1
Gaboon (1864)	19·3	19·8	Gaboon (1864)	10·3	12·9
Gaboon (1880)	17·6	18·5	Gaboon (1880)	10·1	11·9
Mean 19·0			*Mean 11·5*		
Philippines ...	17·5	—	Philippines ...	11·2	—
Borneo... ...	17·4	18·2	Borneo... ...	10·5	9·7
Mean 17·7			*Mean 10·5*		
Veddahs ...	15·9	18·0	Veddahs ...	12·7	11·9
Hindoos ...	15·2	—	Hindoos ...	11·7	—
Mean 16·4			*Mean 12·1*		
Egyptians ...	13·2	14·6	Egyptians ...	8·7	11·5
English ...	13·7	13·9	English ...	9·3	10·3
Mean 13·8			*Mean 10·0*		
General Means	*17·9*	*18·7*	*General Means*	*12·5*	*12·9*

* Nubians excepted: these were included at a later date.
† Hylobates not included.

Whether we take the number of cases in which the female is more variable than the male—i.e. 7 out of 9 in the first series and 6 out of 10 in the second—or the mean variabilities of the whole series, we see that the female is certainly not less variable than the male.

Massed as above the table seems to show more racial relationship to variability. The races group better together, and generally speaking—Veddahs, perhaps, excepted—there is decreasing variability as we reach higher degrees of civilisation, i.e. as we pass from the Apes through East Asiatics and Negroes to Hindoo, Ancient Egyptian and English. It is as well to see if these results correspond with those for the two mesodacryal and the simotic indices.

It will be seen in the first place that these results (Table XIX) confirm those of the previous table (XVIII) for the variation of the absolute lengths, in that

TABLE XIX.

Average Values of Variability of Indices of Nasal Bridge (α, β and S)

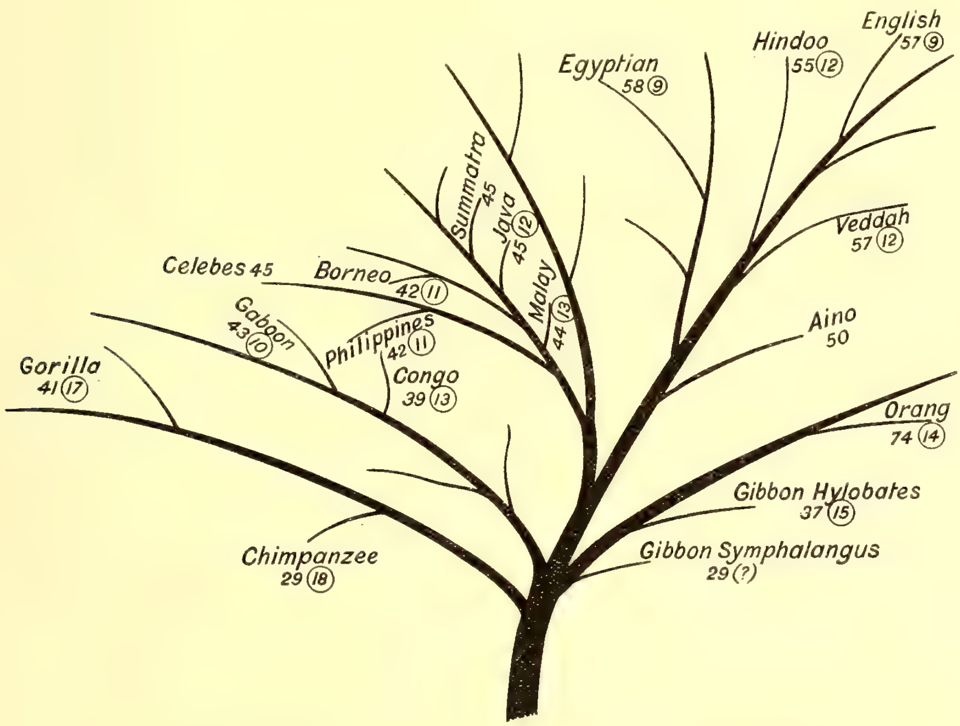
From Standard Deviations			From Coefficients of Variation		
	♂	♀		♂	♀
Chimpanzees ...	6·8	7·8	*Chimpanzees* ...	29·2	35·3
Gorillas ...	14·5	14·4	*Gorillas* ...	31·5	29·8
Orangs ...	30·8	24·0	*Orangs* ...	30·0	21·1
[*Hylobates* ♂ + ♀ 9·9]			[*Hylobates* ♂ + ♀ 30·4]		
*Mean 16·4**			*Mean 29·5**		
Malays ...	8·2	—	Malays ...	20·8	—
Javanese ...	11·2	—	Javanese ...	27·5	—
Mean 9·7			*Mean 24·2*		
Congo	7·3	7·2	Congo	22·9	24·3
Gaboon (1864)	7·4	9·1	Gaboon (1864)	20·4	24·8
Gaboon (1880)	9·3	9·4	Gaboon (1880)	25·0	28·3
Mean 8·3			*Mean 24·3*		
Philippines ...	7·0	—	Philippines ...	19·2	—
Borneo... ...	10·0	8·4	Borneo... ...	28·5	21·9
Mean 8·5			*Mean 23·2*		
Veddahs ...	10·3	14·9	Veddahs ...	18·4	27·7
Hindoos ...	9·7	—	Hindoos ...	18·5	—
Mean 11·6			*Mean 21·5*		
Egyptians ...	9·3	8·9	Egyptians ...	16·5	18·6
English ...	10·6	8·3	English ...	18·9	16·3
Mean 9·3			*Mean 17·6*		
General Means	*10·9*	*11·2*	*General Means*	*23·4*	*24·8*

* Hylobates not included.

they make the female slightly more variable on the average than the male. For the standard deviations the male is more variable in 6 out of 10 cases; for the coefficients of variation the female is more variable in 6 out of 10 cases. In either method of measurement the apes are more variable than men, with the one exception of the Chimpanzees in the case of the standard deviations, where the small variation is associated with small variation in the Negro group.

(8) *Racial Relationship as based on Nasal Bridge Measurements.*

Having general regard to Tables XV—XIX for variability, and comparing the orders therein with those for the racial order of absolute lengths, it seems impossible to suggest any *linear* scale of arrangement which will mark racial relationship; it is impossible to assert that the Negroes or the East Asiatic group stand regularly higher or lower in a *linear* scale. Still there does appear to be some order in the scales both for absolute mesodacryal values and for variabilities. Thus the Veddah nasal bridge is more closely related to that of the Orang than the latter's to the Negro group, which on the whole is closer to that of the African apes. Our numbers are too small, our probable errors too large, for insistance on any individual point, but the measurements lead to constants both for index and variability suggesting a scheme of the following type:



Tree. Mesodacryal Subtense Index and Average Variability for ♂'s indicated.

On this tree we have placed the mean values of the mesodacryal subtense index (α) and the average variability of the mesodacryal lengths*, both for males to indicate the degree of correspondence suggested.

In such a tree as this we see the general tendency of the higher races to have (1) a more developed nasal bridge, and (2) a markedly lessened variability. The tree is merely diagrammatic and no great stress is of course laid on it, but it does suggest that the results reached both for mean values and variabilities may not be merely contradictory, but correspond to a real evolutionary order in the development of the nasal bridge.

(9) *On the Correlations of the Nasal Bridge Characters.*

The correlations of the two mesodacryal indices and the simotic index are given in Table XX.

TABLE XX.

Coefficients of Correlation. Nasal Bridge Indices.

	$r_{\alpha\beta}$				$r_{\alpha\alpha'}$				$r_{S\alpha}$			
	♂		♀		♂		♀		♂		♀	
Congo	*50*	·881 ± ·021	*27*	·864 ± ·033	*50*	·882 ± ·021	*27*	·867 ± ·032	*50*	·613 ± ·060	*25*	·704 ± ·068
Gaboon, 1864	*48*	·904 ± ·018	*37*	·941 ± ·013	*48*	·905 ± ·018	*37*	·945 ± ·012	*47*	·258 ± ·092	*35*	·509 ± ·084
Gaboon, 1880	*16*	·976 ± ·008	*18*	·959 ± ·013	*16*	·978 ± ·007	*18*	·959 ± ·013	*16*	·423 ± ·138	*18*	·508 ± ·118
Philippine Islands ...	*17*	·725 ± ·078		—	*17*	·727 ± ·077		—	*17*	·374 ± ·141		—
Borneo	*26*	·948 ± ·013		—	*26*	·945 ± ·014		—	*26*	·426 ± ·108		—
Java and Madura ...	*39*	·940 ± ·013		—	*39*	·944 ± ·012		—	*39*	·318 ± ·097		—
Mixed "Malays" ...		—	*24*	·897 ± ·027		—	*24*	·897 ± ·027		—	*23*	·696 ± ·073
Hindoos	*10*	·929 ± ·029		—	*10*	·931 ± ·028		—	*10*	·891 ± ·044		—
English	*50*	·955 ± ·008	*50*	·961 ± ·007	*50*	·956 ± ·008	*50*	·962 ± ·007	*50*	·440 ± ·077	*50*	·519 ± ·070
Egyptians	*50*	·976 ± ·005	*50*	·964 ± ·007	*50*	·975 ± ·005	*50*	·964 ± ·007	*50*	·444 ± ·077	*50*	·471 ± ·074
Veddahs	*15*	·959 ± ·014	*12*	·971 ± ·011	*15*	·958 ± ·014	*12*	·972 ± ·011	*15*	·291 ± ·159	*12*	·369 ± ·168
Chimpanzees	*33*	·675 ± ·064	*18*	·782 ± ·062	*33*	·674 ± ·064	*18*	·784 ± ·061	*13*	·062 ± ·186		—
Gorillas	*27*	·889 ± ·027	*25*	·928 ± ·019	*27*	·888 ± ·027	*25*	·924 ± ·020	*21*	·740 ± ·067	*19*	·464 ± ·121
Orang-utans	*44*	·839 ± ·030	*13*	·720 ± ·090	*44*	·844 ± ·029	*13*	·715 ± ·091		—		—
Gibbons, Hylobates ♂ + ♀	*25*	·812 ± ·046			*25*	·815 ± ·045			*16*	·199 ± ·162		

We have already (see pp. 399—401) discussed the relation of α to α' and seen that it does not lead us to any very marked racial order. The correlation of the two mesodacryal indices gives a somewhat clearer result, if we group the correlations thus:

* Coefficients of variation of arc, chord and subtense.

$r_{\alpha\beta}$ in Groups.

	♂	♀
Apes	·801	·810
East Asiatics ...	·871	·897
Negroes	·920	·931
Veddahs	·959	·971
Hindoos	·929	—
Egyptians } English } ...	·965	·962
All values, means	·892	·899

The suggestion is that there is increased correlation with higher civilisation, the Veddahs being somewhat out of rule. The females are slightly, but very probably not significantly, more highly correlated than the males. There is nothing out of keeping in this correlation order with that for the means and variations exhibited in the "tree" on p. 417, but it is too slender to give it much support. If we turn to the correlations of simotic and mesodacryal (subtense) indices r_{Sa} we find a greater range of variation and thus more possibility of getting an order of some racial interest.

r_{Sa}. *Correlation Order.*

	♂	♀
Chimpanzees ...	·06	—
[*Hylobates* ...	·20 (♂ + ♀)	—]
Gaboon, 1864 ...	·26	·51
Veddahs	·29	·37
Javanese... ...	·32	—
Philippines ...	·37	—
Gaboon, 1880 ...	·42	·51
Borneo	·43	—
Malays	—	·70
English	·44	·52
Egyptians ...	·44	·47
Congo	·61	·70
Gorillas	·74	·46
Means	·40	·53

With the single exception of the Gorilla the females have in every case, and often markedly, more correlation than the males. But the results are extremely irregular, even if we allow for the small numbers dealt with. Accordingly it seemed worth while to deal more at length with the relation between the simotic and mesodacryal characters, and further correlations were worked out. It must

be borne in mind that the mesodacryal subtense consists of two portions, the first part due to the nasal bones is the simotic subtense, the second part is due to the maxillary bones. If we subtract the simotic subtense from the mesodacryal subtense $(DS)-(SS)$, we have a measure of the part of the nose due solely to the maxillary bones, and we may correlate this with (SS) the part due to the nasal bones. There will be no spurious correlation in this as in correlating (DS) and (SS) directly. Further we may form an index $\{(DS)-(SS)\}/(DC)$ and correlate this with the simotic index $(SS)/(SC)$, although in doing this (SC) must be considered in a certain sense as contained in (DC). The spurious correlation, however, is unlikely to be as great in this as in (SS) and (DS), for the maxillary walls of the nose can approach so close that there is no (SC) at all, as in the case of many Orangs and in some Negro skulls*.

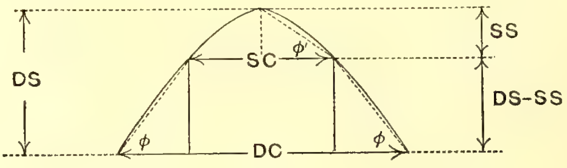


Fig. (i).

(10) *The Maxillary and Simotic Nasal Angles, ϕ and ϕ'.*

The value of $DC-SC$ may be used to obtain a measure of the flatness of the maxillary walls of the nose. Clearly (see figure above)

$$\tan\phi = (DS-SS)/\tfrac{1}{2}(DC-SC).$$

We term ϕ the maxillary nasal angle and its complement $90^\circ-\phi$ measures approximately the angle the maxillary wall of the nose makes with the median plane of the skull. Table XXI records the values of the nasal maxillary angle ϕ for the races dealt with; they are obtained not from the calculation of ϕ from individual crania, but by the rougher process of using the mean values of the mesodacryal and simotic subtenses and chord in the above formula†. The table emphasises again the relation of the Veddahs and the higher races to Orang and Hylobates, while it indicates the relation of the Negro races to the African anthropoids. At some points the female order is not in accord with the male, but the smallness of many of the series dealt with seems sufficient to explain the source of these differences. Some races have clearly a marked sexual difference, but it is not always one way and the means show no very great difference; more data and longer series would be needful to be certain of a persistent sexual differentiation.

* E.g. Gaboon 1864, Nos. 10, 11, 81, and Congo 70: see *Biometrika*, Vol. VIII. p. 319, and see especially the remarks in Tables of measurements. In one gibbon the simotic subtense was actually negative or the nasal bones depressed.

† Table XXIV A gives the values of ϕ and their variabilities for a small number of selected races. It will be seen that the differences due to the two methods are not great. The short method will give results close enough for any conclusions which are in the least likely to be drawn from such small series.

TABLE XXI.

Maxillary Nasal Angle ϕ.

♂		♀	
Orang-utans	55·3°	*Orang-utans*	57·5°
Egyptians	54·3°	Veddahs	56·1°
Veddahs	53·8°	Egyptians	54·3°
[*Hylobates* (♂ + ♀)	52·5°]	[*Hylobates* (♂ + ♀)	52·5°]
Hindoos	52·0°	Gaboon, 1880	50·3°
Nubians	51·3°	English	49·3°
Sumatra	51·3°	Philippines	49·0°
English	50·8°	—	—
Ainos	47·6°	—	—
Moluccas	47·0°	—	—
Celebes	46·9°	Gaboon, 1864	48·5°
Javanese	46·0°	—	—
Gaboon, 1864	45·7°	Hindoos	48·4°
Gaboon, 1880	45·2°	Nubians	48·2°
Malays	44·7°	—	—
Philippines	44·6°	Javanese	44·8°
Borneo	44·4°	Congo	42·5°
Congo	44·1°	[*Symphalangus*	42·3°]
[*Symphalangus* (♂ + ♀)	42·3°]	Borneo	42·0°
Gorillas	35·1°	*Gorillas*	39·1°
Chimpanzees	34·8°	*Chimpanzees*	37·8°
Mean	47·1°	Mean	47·7°

The smaller ϕ, the flatter the maxillary part of nasal bridge.

If we now turn to Table XXII we find a number of additional correlations, which are sufficiently interesting to make one appreciate how much more valuable they would have been had we had longer series and more variety of races available. We have first correlated the mesodacryal and simotic portions, $DS - SS$ and SS, of the whole subtense of the nasal bridge. In every case dealt with, male and female, these subtenses are *negatively* correlated. On the other hand the correlation of the mesodacryal subtense with the simotic subtense is always positive and often substantial*. This positive correlation is therefore due to the appearance of SS in both variates, i.e. it is a spurious correlation. Actually when the maxillary part of the nasal bridge is larger, then the simotic part is smaller. Intraracially—that is among individuals of the same race—there is a compensatory growth of the nasal bridge and those who develop large nasal bones have a smaller maxillary nasal development. This point seems to us of very great importance; the tendency now-a-days is to insist on the importance of anatomical unities. This has largely arisen from two sources. In the first place craniology

* For the mesodacryal subtense with the simotic subtense, all the correlations except that for the male Chimpanzee are positive and often substantial, the intraracial correlation amounting to + ·34 to + ·38. Further the interracial correlation is high indeed, being + ·81 to + ·87. But as we see these results are in part spurious.

TABLE XXII.

Additional Correlations. Nasal Bridge Characters.

	$r_{DS,\ SS}$		$r_{DS-SS,\ SS}$		$r_{\frac{DS-SS}{DC},\ \frac{SS}{SC}}$		$r_{DC,\ SC}$		$r_{DC-SC,\ SC}$	
	♂	♀	♂	♀	♂	♀	♂	♀	♂	♀
Congo	+ ·467 ± ·075	+ ·632 ± ·081	− ·143 ± ·093	− ·138 ± ·109	− ·074 ± ·095	+ ·092 ± ·134	+ ·275 ± ·088	+ ·217 ± ·129	− ·501 ± ·071	− ·729 ± ·063
Gaboon, 1864 ...	+ ·295 ± ·090	+ ·437 ± ·092	− ·487 ± ·075	− ·207 ± ·109	− ·229 ± ·093	+ ·026 ± ·114	+ ·345 ± ·087	+ ·536 ± ·081	− ·745 ± ·044	− ·387 ± ·097
Java and Madura	+ ·658 ± ·061	—	− ·035 ± ·108	—	− ·165 ± ·105	—	+ ·266 ± ·100	—	− ·589 ± ·071	—
English	+ ·355 ± ·083	+ ·505 ± ·071	− ·480 ± ·073	− ·130 ± ·094	− ·102 ± ·094	+ ·274 ± ·088	+ ·446 ± ·076	+ ·611 ± ·060	− ·386 ± ·081	− ·328 ± ·085
Egyptians ...	+ ·416 ± ·079	+ ·230 ± ·090	− ·457 ± ·075	− ·465 ± ·075	− ·132 ± ·094	− ·169 ± ·093	+ ·172 ± ·093	+ ·520 ± ·070	− ·571 ± ·064	− ·138 ± ·094
Gorillas	+ ·327 ± ·131	+ ·120 ± ·153	− ·138 ± ·144	− ·468 ± ·121	+ ·624 ± ·090	− ·079 ± ·154	+ ·512 ± ·109	+ ·303 ± ·141	− ·107 ± ·146	− ·137 ± ·152
Chimpanzees ...	− ·153 ± ·183	+ ·364 ± ·195	− ·587 ± ·123	− ·042 ± ·224	− ·377 ± ·161	− ·079 ± ·223	− ·039 ± ·187	+ ·425 ± ·184	− ·613 ± ·117	+ ·222 ± ·214
[*Hylobates* ...	+ ·116 ± ·166	for ♂ + ♀	− ·220 ± ·161	for ♂ + ♀	− ·112 ± ·166	for ♂ + ♀	− ·195 ± ·162	for ♂ + ♀	− ·730 ± ·079	for ♂ + ♀]
Interracial (omitting *Orangs*)	+ ·868 ± ·039	+ ·806 ± ·066	+ ·581 ± ·105	+ ·501 ± ·140	+ ·370 ± ·137	+ ·393 ± ·158	+ ·344 ± ·140	+ ·612 ± ·117	− ·413 ± ·132	− ·332 ± ·167
Intraracial mean*	+ ·338	+ ·381	− ·333	− ·226	− ·065	+ ·011	+ ·283	+ ·435	− ·502	− ·250

* *Hylobates* excluded.

has been chiefly in the hands of the pure anatomist rather than in those of the student of philosophical evolution. The former has insisted on the importance of growth centres and anatomical unity of measurement; he has usually no appreciation of the relatively high correlations of the parts of the bony structure. He is thus liable to overlook the fact that from the standpoint of evolution a complex anatomical organ may be of far more importance to the race and to the individual than one factor of it which may be an anatomical unity. In the second place Mendelism has become the mode, and to surmount difficulties about characters which do not "mendelise," it has been customary to assert that they consist of a complex of simple Mendelian units*. On the original Mendelian theory such "units" were asserted to be independent, although as such a theory was found to be unworkable, a crude theory of "coupling," as a measurable association of anatomically or physiologically simple unities, was evolved to cover the real facts of correlation. In this manner there has arisen a conscious or unconscious association of anatomical and Mendelian units, and it has been supposed that an anatomical unity would be more likely to "mendelise" than a combination of such unities—a single bone like the femur than a measurement like the leg length or the stature. The present result seems an illustration of the futility of neglecting the correlation of characters or treating them as compounded of independent Mendelian or even anatomical unities. The bridge of the nose consists of parts which are distinct anatomical unities, but the evolutionary factor is probably far better represented by the physiognomic factor—the whole bridge of the nose—than by any anatomical unit. When one part exceeds there is a correlated defect in the other part, and this rule extends practically through all the races examined. It is thus quite conceivable that a single measurement like the sagittal arc is from the evolutionary standpoint far more important than any of the simple arcs—opisthion to lambda, lambda to bregma, and bregma to nasion—into which we may reduce it by aiming at the measurement of anatomical unities. There is such a thing as an evolutionary factor, a determinant which controls the development of a whole series of anatomical unities, and the heredity brought into evidence by such a factor may be far more important from the standpoint of evolution than the heredity of any simple anatomical character. What group of anatomical units go to form any such determinant can only be appreciated by a thorough study of the correlations of simple anatomical parts.

Especially in the nasal bridge is this matter of very great importance, for while within the race an individual with small maxillary height ($DS - SS$) has a large simotic height (SS), *interracially* this is no longer true, the race with a small maxillary height will have a small simotic height and the positive correlation interracially is very substantial ($+ \cdot 581$)—this must be compared with the mean intraracial correlation which is $- \cdot 333$. Both these results are for males, for females the numbers are $+ \cdot 501$ and $- \cdot 226$ respectively. This reversal of sign

* Of course when it is convenient a vast congeries of factors, such as "Albinism" or "Jewishness," are illogically treated as a simple Mendelian "unit"!

when we pass from intraracial to interracial correlations is a thing previously not unknown. Thus in 1903 in a paper* on interracial and intraracial correlations, Jacob, Lee and Pearson showed that while with individuals a long cranium is likely to be a broad cranium, yet interracially a compensating factor comes in, a long-headed race is likely to be a narrow-headed race. Thus it is quite conceivable that capacity of the cranium—depending on many anatomical units—may be of far more evolutionary importance than the measure of any single "anatomical unit" of the skull. We have referred at length to these matters here, because there is a growing idea—fostered probably by the idea of Mendelian "units," that the measurement of "anatomical units"—or the measurement between "anatomical points"—is of primary importance in craniology. Whether craniologists measure the *same* thing or not is immensely important, the correlation between anatomical units is also of great value as determining what combinations of simple units form evolutionary factors. But very little service is done by insisting too largely on anatomical unity in and for itself. Evolution depends largely on physiological fitness, and organs of physiological importance are rarely compounded of either single anatomical or of single Mendelian units.

If we now turn to the index correlations of $(DS - SS)/DC$—i.e. height of maxillary portion of nose by its breadth—with SS/SC—i.e. height of nasal bones portion by its breadth—our results are less uniform. Putting the male Gorillas on one side for a moment there are only two of the correlations which can be said to be significant, namely, possibly the male Chimpanzee and the female English, and these are of opposite signs. The means for male races and for female races are $-\cdot 065$ and $+\cdot 011$, and these would be of no service for prediction. With regard to the Gorilla result, this arises principally from three males with the very high simotic indices of 88, 106 and 103. It is conceivable that the material is not really homogeneous; the arithmetic has been carefully repeated without change in values. If we take these indices to measure the cuneal or wedge-shaped properties of the two portions of the nose, we should say that individually there is very little relation between the ratio of height to breadth in the simotic and maxillary portions of the nasal bridge; the only exception to this rule being the male Gorilla. On the other hand the cuneal characters are *positively* and sensibly correlated $(+\cdot 370$ to $+\cdot 393)$ interracially, an obtuso-cuneal nasal part being also associated with an obtuso-cuneal maxillary part and an acuto-cuneal nasal part with an acuto-cuneal maxillary part.

We have already discussed the maxillary nasal angle ϕ, the complement of which roughly measures the angle between the maxillary wall of the nose at the bridge and the sagittal plane of the skull. We can introduce a similar angle ϕ' to measure roughly the angle between the nasal bone and the simotic chord. We may take (see Fig. (i), p. 420):

$$\begin{aligned}\tan \phi' &= SS/(\tfrac{1}{2} SC) \\ &= 2 \text{ Simotic Index.}\end{aligned}$$

* *Biometrika*, Vol. II. p. 355. For the first introduction of the ideas of *intraracial* and *interracial* correlation, see *Biometrika*, Vol. I. p. 460.

It will be of interest to investigate whether the correlations of ϕ and ϕ' are higher than those of the cuneal indices. The following table* (XXIII) gives the

TABLE XXIII. *Simotic Nasal Angle* ϕ'.

♂		♀	
Gorillas	48·6°	*Gorillas*	47·8°
English	45·5°	English	43·0°
Hindoos	41·8°	Hindoos	38·0°
Egyptians	41·6°	Egyptians	36·5°
Veddahs	41·1°	Veddahs	36·4°
Ainos	40·9°	—	—
Nubians	34·3°	Borneo	34·8°
Javanese	34·1°	Javanese	33·2°
Malays	34·0°	—	—
Gaboon, 1864	31·7°	Philippines	30·1°
Philippines	31·3°	Gaboon, 1864	29·0°
Borneo	31·1°	Congo	27·2°
Celebes	30·7°	—	—
Sumatra	30·3°	—	—
Gaboon, 1880	30·1°	Nubians	25·2°
Moluccas	28·0°	—	—
Congo	27·1°	Gaboon, 1880	24·6°
Chimpanzees	23·2°	*Chimpanzees*	20·3°
†*Symphalangus*	11·5°	†*Symphalangus*	11·5°
†*Hylobates*	9·8°	†*Hylobates*	9·8°
Mean	32·3°	Mean	29·8°

The smaller ϕ' the flatter the simotic portion of the nasal bridge, i.e. the greater the simosis.

TABLE XXIV A. *Maxillary Nasal Angle* ϕ. *Mean and Variability.*

	Male			Female		
	Mean	Standard Deviation	Coefficient of Variation	Mean	Standard Deviation	Coefficient of Variation
Egyptians ...	54·33 ± ·57	5·97 ± ·40	10·99 ± ·75	54·29 ± ·53	5·56 ± ·37	10·24 ± ·70
[*Hylobates* (♂ + ♀)	53·06 ± 1·50	8·90 ± 1·06	16·78 ± 2·06]	—	—	—
English	50·63 ± ·55	5·74 ± ·39	11·35 ± ·78	49·15 ± ·48	5·00 ± ·34	10·17 ± ·69
Javanese	46·06 ± ·91	8·43 ± ·64	18·31 ± 1·44	—	—	—
Gaboon, 1864 ...	45·94 ± ·53	5·34 ± ·37	11·63 ± ·82	49·19 ± ·64	5·62 ± ·45	11·43 ± ·93
Congo	44·13 ± ·53	5·52 ± ·37	12·50 ± ·86	43·30 ± ·71	5·27 ± ·50	12·16 ± 1·18
Gorillas	36·12 ± ·96	6·56 ± ·68	18·15 ± 1·95	39·92 ± 1·10	7·11 ± ·78	17·82 ± 2·01
Chimpanzees ...	35·96 ± 1·80	9·61 ± 1·27	26·74 ± 3·78	34·39 ± 1·82	8·12 ± 1·29	23·60 ± 3·96

* These values of ϕ' are calculated from the mean racial simotic index. In Table XXIV B the mean value of ϕ' as obtained from averaging the simotic angles of the individual members of the race is given for a few races as well as the variabilities for these selected races. The agreement of the directly and indirectly calculated means is quite reasonable.

† Both sexes together.

interracial values of ϕ'. It will be seen of course that the order must be the same as that of the simotic index, but if we compare this table with Table XXI some interesting points arise as we shall show in the next section. Tables XXIV A and XXIV B give the means and variabilities of the maxillary and simotic angles ϕ and ϕ'.

TABLE XXIV B.

Simotic Nasal Angle ϕ'. Mean and Variability.

	Male			Female		
	Mean	Standard Deviation	Coefficient of Variation	Mean	Standard Deviation	Coefficient of Variation
Egyptians ...	41·25 ± ·45	4·70 ± ·32	11·40 ± ·78	35·97 ± ·57	6·02 ± ·41	16·73 ± 1·16
[*Hylobates* (♂ + ♀)	9·75 ± ·84	4·97 ± ·59	50·99 ± 7·50]	—	—	—
English	44·70 ± ·63	6·58 ± ·44	14·73 ± 1·01	42·58 ± ·47	4·96 ± ·33	11·65 ± ·80
Javanese	33·23 ± ·88	8·15 ± ·62	24·54 ± 1·98	—	—	—
Gaboon, 1864 ...	31·15 ± ·72	7·34 ± ·51	23·56 ± 1·73	28·46 ± ·85	7·42 ± ·60	26·08 ± 2·24
Congo	26·62 ± ·73	7·68 ± ·52	28·86 ± 2·10	26·48 ± 1·20	8·91 ± ·85	33·63 ± 3·55
Gorillas	46·48 ± 1·52	10·30 ± 1·07	22·16 ± 2·42	45·63 ± 1·74	11·23 ± 1·23	24·61 ± 2·85
Chimpanzees ...	22·77 ± 1·37	7·31 ± ·97	32·10 ± 4·66	20·00 ± 1·73	7·70 ± 1·22	38·51 ± 6·97

It would appear from Tables XXI and XXIII that ϕ' has a greater interracial range and from Tables XXIV A and B that it has a greater intraracial variability than ϕ. What we need for a good racial character is a great interracial range and a small intraracial variability. We cannot therefore say that ϕ' is a better racial character than ϕ.

(11) *General Classification of Nasal Bridges by ϕ and ϕ'**.

We need some new terms. We propose to term races which have high values or low values of both ϕ and ϕ' *homoclinic*, while those which have low values of one and high of the other are *heteroclinic*. Among apes the Gorilla is markedly heteroclinic, and Hylobates is the same, but in the opposite direction. The Chimpanzee and Symphalangus are both homoclinic and so are the English, Egyptians, Hindoos and Veddahs.

Without regard to the exact numerical values and as merely descriptive terms the following types of Gephyrosis will indicate the nature of the characterisation, where it must be remembered that the reader has to fix his attention on the

* Professor G. D. Thane has kindly made the following suggestions for a more complete angle nomenclature of the nasal bridge (see Fig. (i), p. 420):

Lateral simotic angle = our simotic nasal angle ϕ'.

Mesial simotic angle = our $90° - \phi'$ = angle between simotic chord and median plane.

Simotic angle (median) = our $180° - 2\phi'$ = internal angle between two simotic chords.

Lateral maxillary angle = our ϕ = maxillary nasal angle.

Maxilli-simotic angle = our $180° - \phi + \phi'$ = internal angle between simotic and maxillary chords.

The most important point here is the maxilli-simotic angle, the approach of which to 180°, i.e. the

lower parts of the "bridge" from its "springings" at the dacrya to the strokes marking the naso-maxillary sutures.

14	23	27	34	22
Hyperplatygephyreal	*Platygephyreal*	*Mesogephyreal*	*Leptogephyreal*	*Hypoleptogephyreal* (approaching *Orthogephyreal*)

FIG. (ii). *Types of Gephyrosis.*

Similarly we have for the part of the "bridge" from "haunch" to "haunch," i.e. the nasal bones, or part between the strokes marking the naso-maxillary sutures, the following characterisation:

15	8	10	29	13
Hypersimotic	*Simotic*	*Mesotic*	*Stenotic*	*Hypostenotic*

FIG. (iii). *Types of Simosis.*

For classificatory purposes it may be useful to divide ϕ into three classes: when ϕ is under 46°, the group shall be called *platygephyreal*, when ϕ lies between 46° and 51°, the group is *mesogephyreal*, and when ϕ is over 51°, the

equality of ϕ and ϕ', marks the "triangularity" of the bridge, or the unbroken character of the nasal wall.

Maxilli-simotic angle $180° - \phi + \phi'$.

♂		♀	
Gorillas	193·5°	*Gorillas*	188·7°
English	174·7°	English	173·7°
Ainos	173·3°	—	—
Hindoos	169·8°	Borneo	172·8°
Malays	169·3°	—	—
Chimpanzees	168·4°	Hindoos	169·6°
Javanese	168·1°	Javanese	168·4°
Veddahs	167·3°	Congo	164·7°
Egyptians	167·3°	*Chimpanzees*	162·5°
Borneo	166·7°	Egyptians	162·2°
Philippines	166·7°	Philippines	161·1°
Gaboon (1864)	166·0°	Gaboon (1864)	160·5°
Gaboon (1880)	164·9°	Veddahs	160·3°
Celebes	163·8°	—	—
Congo	163·0°	Nubians	157·0°
Nubians	163·0°	Gaboon (1880)	154·3°
Moluccas	161·0°	—	—
Sumatra	159·0°	—	—
Symphalangus	149·2°	*Symphalangus*	149·2°
Hylobates	137·3°	*Hylobates*	137·3°

Only the Gorillas have a re-entrant racial maxilli-simotic angle, although such angles occur in individual cases of other races. Of the races considered the English appear to have most nearly an unbroken nasal wall. The racial order is not clearly marked as Negroes and Asiatics are much intermixed.

TYPES OF THE NASAL BRIDGE.

1 **2** **1** *bis* **2** *bis*

Dog. Nasal Bridge. Dog. Muzzle Section.

1—2 *bis* = Hypersimotic Platygephyrosis.

3 Chimpanzee. **4** Chimpanzee. **5** Symphalangus.

6 **7** **8** **9**

Congo Negroes. Gaboon Negroes.

3—9 = Simotic Platygephyrosis.

10 Malay. **11** Malay. **12** Malay.

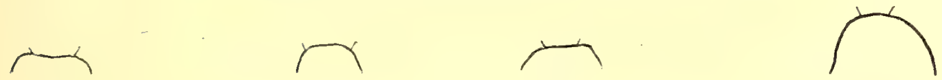
10—12 = Mesotic Platygephyrosis.

13 Gorilla. **14** Gorilla.

13—14 = Stenotic Platygephyrosis.

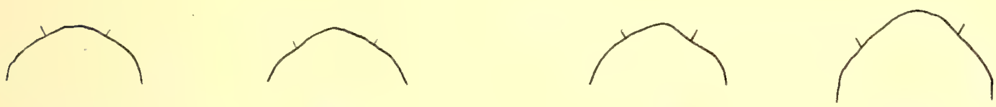
It has not always been possible to select a nasal bridge with the mean values of both ϕ and ϕ', and the typical form will hardly be realised even from two or three cases.

TYPES OF THE NASAL BRIDGE.



15 Hylobates. 16 Hylobates. 17 Hylobates. 18 Orang-utan.

15 = Hypersimotic Mesogephyrosis. 16—17 = Simotic Leptogephyrosis. 18 = Simotic Leptogephyrosis.

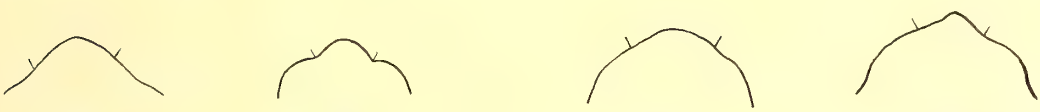


19 20 Natives of Moluccas.

19—20 = Simotic Mesogephyrosis.

21 22 Natives of Sumatra.

21—22 = Simotic Leptogephyrosis.



23 Aino. 24 Aino (Variant).

23—24 = Mesotic Mesogephyrosis.

25 Nubian. 26 Nubian (Variant).

25—26 = Mesotic Leptogephyrosis.

27 English. 29 Veddah. 31 Egyptian. 33 Hindoo.

28 English. 30 Veddah. 32 Egyptian. 34 Hindoo.

27—28 = Stenotic Mesogephyrosis. 29—34 = Stenotic Leptogephyrosis.

It has not always been possible to select a nasal bridge with the mean values of both ϕ and ϕ', and the typical form will hardly be realised even from two or three cases.

group is *leptogephyreal*. In the same way the simotic nasal angle ϕ' may be classed as *simotic* when under 32°, *mesotic* between 32° and 41°, and *stenotic* when over 41°. These limits should all be reduced by about 2° in classifying the female simotic nasal angle, which differs sexually by at least this amount from the male. We have accordingly the following table, of which we give illustrations to indicate the type. The diagonal from simotic platygephyrosis to

TABLE XXV.

Value of Maxillary Nasal Angle ϕ.

Value of Simotic Nasal Angle ϕ'.	Under *46°*	*46°* to *51°*	Over *51°*
Under *32°*	Simotic Platygephyrosis e.g. *Chimpanzees* and Negroes	Simotic Mesogephyrosis e.g. Natives of Moluccas	Simotic Leptogephyrosis e.g. *Hylobates* and natives of Sumatra
32° to *41°*	Mesotic Platygephyrosis e.g. Malays	Mesotic Mesogephyrosis e.g. Ainos	Mesotic Leptogephyrosis e.g. Nubians
Over *41°*	Stenotic Platygephyrosis e.g. *Gorillas*	Stenotic Mesogephyrosis e.g. English	Stenotic Leptogephyrosis e.g. Egyptians and Hindoos

stenotic plagygephyrosis marks the homoclinic groups. The simotic leptogephyrosis of the smaller gibbon and the stenotic platygephyrosis of the gorilla mark the heteroclinic extremes, which weaken the interracial correlation of ϕ and ϕ'. Table XXVI gives the values of $r_{\phi\phi'}$.

With the exception of the Gorilla, which is as in the case of the cuneal indices (p. 424) again anomalous, we see that there is *intraracially* a small *negative* correlation between flatness of the maxillary and flatness of the nasal bones. But *interracially* it is otherwise; there is quite a moderate degree of positive correlation and a flat nasal part of the bridge is found in the same races as a flat maxillary part. This is another instance of the reversal of intraracial by interracial correlation.

(12) *Discussion of Projections of Nasal and Maxillary Portions of Nose on the Base of the Nasal Bridge.*

Having studied the relation of the subtenses, the indices and the angles, we now turn to the chords. The mesodacryal chord is positively correlated with the simotic chord for all races except male Chimpanzees and the mixed data for Hylobates, but in both these cases the correlations are insignificant having regard to their probable errors; see Table XXII, p. 422. Both interracially and intraracially a broad maxillary

TABLE XXVI.

Correlation of Maxillary and Simotic Nasal Angles ϕ and ϕ'.

	♂	♀
Congo	$-\cdot 037 \pm \cdot 095$	$-\cdot 127 \pm \cdot 133$
Gaboon, 1864 ...	$-\cdot 338 \pm \cdot 087$	$-\cdot 095 \pm \cdot 113$
Javanese	$-\cdot 054 \pm \cdot 108$	—
English	$-\cdot 169 \pm \cdot 093$	$-\cdot 005 \pm \cdot 095$
Egyptians	$+\cdot 053 \pm \cdot 095$	$-\cdot 183 \pm \cdot 092$
Gorillas	$+\cdot 447 \pm \cdot 118$	$-\cdot 217 \pm \cdot 147$
Chimpanzees	$-\cdot 329 \pm \cdot 167$	$-\cdot 247 \pm \cdot 211$
[*Hylobates* (♂ + ♀) ...	$-\cdot 058 \pm \cdot 168$	$-\cdot 058 \pm \cdot 168$]
Intraracial Mean ...	$-\cdot 061$	$-\cdot 146$
Interracial Value* ...	$+\cdot 353 \pm \cdot 139$	$+\cdot 220 \pm \cdot 178$

portion of the nose would thus signify a broad simotic portion. But the whole result is, as that for the subtenses, spurious and arises from the simotic portion recurring in the mesodacryal chord. If we subtract the simotic chord from the mesodacryal then the difference $(DC - SC)$ is *negatively* correlated with the simotic chord SC, and the correlation is fairly substantial both interracially and intraracially. In other words, whenever a race or an individual is found for which the nasal bones have a big projection on the base of the nasal bridge, then the maxillary bones have a small projection, and *vice versâ*. Thus individually and racially a principle of compensation appears at work. Now we have seen that intraracially ϕ is negatively correlated with ϕ', or a steep maxillary portion is associated in the individual with a flat nasal bone portion of the bridge. Thus intraracially it is conceivable that even if the breadth of the nasal bone and the breadth of the maxillary wall of the nose were positively correlated, the projections might well be of opposite signs. But we cannot use this explanation to account for the negative interracial correlation of the projections, for here the maxillary and simotic nasal angles are positively correlated. In the case of intraracial correlations not only are the projections $\frac{1}{2}SC$ and $\frac{1}{2}(DC - SC)$ negatively correlated but also the subtenses SS and $(DS - SS)$. Thus it is not improbable that in the individual a small nasal bone is associated with a large maxillary wall of the nose—a feature we may speak of as the Orang-utan character. But this does not hold interracially, for the subtenses SS and $(DS - SS)$ are *positively* correlated (Table XXII) and substantially so for both sexes. It is of some interest to consider whether interracially the nasal bone and the maxillary wall of the nose are

* *Hylobates* excluded.

positively correlated. In order to test this point a little more definitely, let us suppose SB to be the breadth of a nasal bone and DB the breadth of the maxillary wall at the bridge. *Approximately*:

$$SB = \sqrt{\tfrac{1}{4}(SC)^2 + (SS)^2} = (SS)/\sin\phi',$$

$$DB = \sqrt{\tfrac{1}{4}(DC - SC)^2 + (DS - SS)^2} = (DS - SS)/\sin\phi.$$

Table XXVII gives the values of SB and DB found for each race from the means of that race for SC, SS, DC and DS. Of course the nasal bone and the

TABLE XXVII.

Calculated Values of SB, the breadth of either nasal bone at the bridge and DB, the breadth of the maxillary portion of the nasal wall at the bridge.*

Race	♂		♀	
	SB	*DB*	*SB*	*DB*
Congo	5·44	9·83	5·37	9·21
Gaboon, 1864	5·48	9·82	5·53	9·53
Gaboon, 1880	4·59	10·54	5·12	9·71
Philippines	4·91	8·36	5·14	8·06
Moluccas	5·28	9·12	—	—
Borneo	4·51	9·33	4·54	8·09
Malays	4·69	9·40	—	—
Celebes	5·05	9·34	—	—
Sumatra	5·71	8·94	—	—
Javanese	5·33	9·23	5·11	7·64
Nubians	6·64	10·22	6·60	8·60
Ainos	5·80	8·96	—	—
Hindoos	5·67	8·73	4·91	8·32
English	6·56	9·56	6·10	8·77
Egyptians	6·97	9·49	6·67	8·39
Veddahs	5·98	9·29	5·19	9·17
Chimpanzees	4·52	6·41	4·24	5·91
Gorillas	4·41	9·97	3·81	9·61
Orang-utans	—	[13·57]	—	[11·02]
Hylobates (♂ + ♀) ...	2·70	4·27	2·70	4·27
Symphalangus (♂ + ♀) ...	3·11	4·86	3·11	4·86

maxillary wall are often, as our contours (pp. 428—9) show, considerably curved, but the chord of these arcs will indicate, at any rate approximately, what is the nature of the interracial correlations. Omitting Hylobates and Symphalangus as unsexed and the Orangs we find:

Interracial correlation of nasal bone and maxillary wall

Males $+ \cdot 238$, Females $+ \cdot 215$.

* These values are really the hypotenuses or chords corresponding to the nasal and maxillary portions of the bridge of the nose, and these portions are often curved.

It appears therefore reasonable to assume that racially, where the nasal portion is large, so also is the maxillary portion of the nose. Notwithstanding this and the positive correlation of the simotic and maxillary nasal angles, as well as of the subtenses, the projections of the nasal and maxillary portions on the base of the nasal bridge are *negatively* correlated. The algebraic explanation of this paradox is given below*, but it seems probable that its physiological explanation lies in the relative constancy of the mesodacryal chord. If the reader will examine Table XIV he will see that the variability of the simotic chord relative to its size is almost twice that of the mesodacryal chord, and although some of this may be due to the relatively greater difficulty of accurate measurement the bulk of it is not. The following results will indicate the relative stability of the mesodacryal chord. They are deduced from Tables VIII and XIII, males.

	Mesodacryal Chord	Simotic Chord
Interracial Mean	21·6	8·47
Mean racial S. D.	2·44	1·97
Interracial S. D.	1·37	1·20
$100 \times \dfrac{\text{Mean racial S. D.}}{\text{Interracial Mean}}$	11·3	23·3
$100 \times \dfrac{\text{Interracial S. D.}}{\text{Interracial Mean}}$	6·34	14·13
$100 \times \dfrac{\text{Racial Range}}{\text{Interracial Mean}}$	23·5	58·9

Whichever method we take to measure the variation we see that the mesodacryal chord is far less variable than the simotic chord. The breadth of the base of the nasal bridge is relatively constant. The nasal bridge has to span the nasal base, and whether we deal with the problem from the interracial or intraracial standpoint we realise that the nasal structure has to be considered as a whole, and that its anatomical units are very far from being necessarily evolutionary units, or in

* The problem is algebraically of the following kind: given two right-angled triangles ABC and $A'B'C'$ with C and C' for right angles, then with the usual notation $c^2 = a^2 + b^2$, $c'^2 = a'^2 + b'^2$, is it possible for c and c', a and a' and A and A' to be positively correlated together and yet b and b' to be negatively correlated? Clearly if the means be denoted in the usual manner, we have approximately:

$$\bar{b}\,\delta b = \bar{c}\,\delta c - \bar{a}\,\delta a, \qquad \bar{b}'\,\delta b' = \bar{c}'\,\delta c' - \bar{a}'\,\delta a'.$$

Hence multiplying together, summing for all possible pairs and dividing by their number, we have

$$\bar{b}\bar{b}'\,\sigma_b\sigma_{b'}\,r_{bb'} = \bar{c}\bar{c}'\,\sigma_c\sigma_{c'}\,r_{cc'} + \bar{a}\bar{a}'\,\sigma_a\sigma_{a'}\,r_{aa'} - \bar{c}\bar{a}'\,\sigma_c\sigma_{a'}\,r_{ca'} - \bar{c}'\bar{a}\,\sigma_{c'}\sigma_a\,r_{c'a}.$$

Hence if $r_{cc'}$ and $r_{aa'}$ are positive as the hypothesis supposes, and at least one or both $r_{ca'}$ and $r_{c'a}$ are positive also, the negative term or terms on the right may exceed the positive and $r_{bb'}$ be *negative*. In the case of the simotic and maxillary nasal triangles both $r_{ca'}$ and $r_{c'a}$ are positive, and the last pair of terms on the right is wholly negative and in excess of the first pair. The above result depends of course on $\bar{c}^2 = \bar{a}^2 + \bar{b}^2$ and $\bar{c}'^2 = \bar{a}'^2 + \bar{b}'^2$ being nearly true. These give $\bar{c}^2 = 84{\cdot}83$ against actual 84·64 and $c'^2 = 29{\cdot}68$ against 29·16. These are close enough to justify the use of the above formula, which on putting in the actual numbers gave $r_{bb'}$ the correlation of the projections negative, as the direct investigation gave it. Thus the origin of the negative value for the projections' correlation lies in the negative terms involving the cross-correlations of nasal bone with maxillary height and simotic subtense with maxillary wall, which are themselves positive.

themselves primarily significant for interracial comparison or evolutionary history. The intimate correlations of anatomical units renders them also very frequently of small importance as hereditary characters, and a complex of anatomical units may be much closer to an evolutionary unit,—that entity which as a whole is the subject of selective action. For respiratory efficiency the mesodacryal arc may be of far more importance than either of its components, the maxillary or nasal bone portions. Thus we think it probable that the value of anatomical units in craniometric measurements is liable to be overrated, and in view of their complex and subtle intercorrelations their identification with *independent* "Mendelian units"—whatever the latter may signify—is unjustifiable.

In the series of cuts on pp. 428—9, we have endeavoured to give "type" contours of the nasal bridge for various races on the basis of the classification adopted in Table XXV. The reader is warned that in any race individuals may be found with many of these types. Our "types" are those of individuals whose simotic and maxillary nasal angles, ϕ' and ϕ, are as near as we could get to the mean values of those angles in the race. Not too much weight must of course be placed on the classification of races thus reached, but the diagrams are not without suggestiveness. We have added the Dog with its *negative* simotic subtense, a condition which is just reached in one or two Gibbons. In the case of the dog, however, the dacryon to dacryon section seems to miss the physiognomic nasal bridge, and we have given also a lower section termed on p. 428 the muzzle section. The short strokes mark for each contour the limits of the nasal bone; the contours are in each case from dacryon to dacryon, and were traced after orientating the skull so that the two dacrya and the summit of the bridge were in one horizontal plane. Obviously larger numbers of many more races would provide a more adequate classification, especially if extended into the animal kingdom generally*.

* The Gibraltar skull is the only "ancient type" of man for which we have been able to obtain the nasal bridge constants. They are:

Mesodacryal	Subtense	8·6,	Simotic	Subtense	2·8
,,	Chord	22·0?	,,	Chord	10·8
,,	Arc	31·0?			

These give

$$\alpha = 39{\cdot}1,\quad \alpha' = 44{\cdot}4,\quad \beta = 40{\cdot}9,\quad S = 25{\cdot}9,\quad \phi = 46^\circ{\cdot}0,\quad \phi' = 27^\circ{\cdot}5,$$

the nasal bridge is thus on the border of the simotic platygephyreal and simotic mesogephyreal groups. The difficulty, however, of determining the dacrya in the absence of the sutures in the cast used is very

Gibraltar Skull.

considerable. The contour of the nasal bridge is shown above. It will be seen to be in no way markedly differentiated. If we accept it as approximating to that of the original, that original would have stood somewhere between the Negroes and the natives of the Moluccas for both the maxillary and simotic nasal angles, i.e. at the Chimpanzee and Symphalangus end of the scales. But the evidence is too fragile to lay any stress on this point.

(13) *Concluding Remarks.* We are aware how much of the present paper is open to criticism on account of the limited number of the racial series dealt with and the sparseness of the individual crania in those series. Notwithstanding we hope that the reader may find suggestive paths leading off from these preliminary researches. We believe that the time must come when in the larger Universities there will exist established schools of craniometry with adequate anatomical, physical and statistical training, and that such schools will carry on secular work in collecting and measuring. We have indicated in this paper that the study of the personal equation of craniometricians has hardly begun, but we look forward to the day when it will be as customary for one of the workers in an established craniometry laboratory to be adopted as a standard and his personal equation relative to other workers at home and abroad discussed, as it is for one astronomical observer to be tested against a second. The intercomparability of measurements is largely taken on blind trust by craniologists to-day. Only when relative personal equation has been studied will it be possible safely to pool the measurements of small series made by different laboratories.

Again, anatomical and zoological museums and collections are singularly defective at the present time. If it were desirable to make a thorough study of the cranium of any animal other than man—say dog or horse—is there any collection which could place at the disposal of the inquirer a hundred crania of definite sex, fairly uniform age and reasonable limits of race? We sadly fear not, and yet how much could be learnt of evolutionary descent by an exact study of a variety of species based on even a hundred crania from each! Take the plane of the foramen magnum, determined say by a plane through basion and opisthion perpendicular to the sagittal plane, and consider the angle this makes with the horizontal plane of the skull, as determined, say, by the Frankfurt Concordat. This angle —closely allied to Daubenton's angle—is of very great evolutionary importance, but who can say—on the basis of really numerically adequate measurements* and other than from mere impression gained on a few specimens—what its value is in the chief mammals, in the anthropoid apes, in negroes and in the various other races of man? Comparative craniometry is almost at its origins if we refer, not to descriptions of "typical"† crania, but to the measurement of numerically adequate series. Not until investigations, only roughly foreshadowed in such a paper as the present, have been made on a far greater variety of species and on a wide range of cranial characters—and this will have to wait until adequate material accumulates in our museums—would any suggestion of evolutionary descent pass into comparative certainty. When the reader realises that some 25 gorilla skulls of one sex, but of all ages

* We do not wish for a moment to underrate Broca's work in this field; but besides his views as to the horizontal planes, his series are often unsexed, and, especially for the chief mammals, wholly inadequate numerically. Thus he uses two or three apes, dogs, cats, and so forth, where we need 50 to 100 of each breed and sex.

† How few anatomists realise that nobody knows what a "typical" cranium is until a *long* series of the particular race has been measured! We remember seeing in one museum a skull, labelled "typical English," and this at a date when not 50 English crania from one district and period had been measured.

and all districts, are about all that the English worker has to rely upon at present, and that it may be doubted whether he could find the same number of dog skulls of known sex, of *one breed* and of fairly certain age, anywhere in the country, he will, we hope, pardon one source of the inadequacy of the present study.

In choosing the nasal bridge we had in view the importance of the nasal organs as a factor of survival, and we believe we have at least indicated that with caution the parts of the bridge can be measured and comparative results obtained. The measurements lead directly, we think, to the suggestion of a racial scale and throw light on the importance of dealing with groups of correlated characters and not with isolated anatomical units. The classification we have suggested, both in the limits of its categories and its terminology, is of course not put forward as final, but we believe it will be helpful as a step to a good ultimate classification and as enabling meanwhile the characters of the nasal bridge of any other race to be associated with those treated in this paper. The needful measurements are simotic chord and simotic subtense, and again mesodacryal chord and subtense; from the first two the simotic nasal angle can be readily found and from all four the maxillary nasal angle. The two taken in conjunction will enable the recorder to classify the nose by aid of Table XXV.

On the whole we consider that Mérejkowsky broke fruitful ground, especially, if the discussion of the nasal bridge be extended to the mesodacryal chord and subtense, which appear of more evolutionary importance than his simotic chord and subtense; but there will undoubtedly have to be standardisation of individual craniometricians in a far higher degree than even for the macrometric measurements of the skull, if these micrometric measurements are to be dealt with confidently.

Addendum. Since this paper was written, it occurred to us that the Moriori crania from Chatham Island in the Royal College of Surgeons might be worth investigating from the standpoint of the nasal bridge—as they are undoubtedly worth investigating from their general craniometric importance.

The nasal bridges, however, turned out to be of no markedly primitive type, as the accompanying contours will indicate.

Moriori Nasal Contours.

They are mesotic mesogephyreal, and stand between Malay and Aino with:

	♂	♀
ϕ	46·0°	49·0°
ϕ'	40·6°	34·3°

48 crania were taken as male, 15 as female.

There is a somewhat noteworthy difference between male and female, but a still more remarkable feature of the results is that for the *whole* eleven characters discussed the male is more variable and in some cases no less than twice as variable as the female. Such exaggerated differences in variation in the sexes are so unusual, especially in primitive races, that we believe that the noteworthy cranial characters of this race, in which "maleness" seems almost a racial character, may have led to a far too stringent selection of female crania. Thus on the present sexing the female coefficient of variation of the mesodacryal subtense is 8·0, of the mesodacryal index α, 8·0, and of the simotic index 14·4,—all values much below any reached for females of the human races given in Table XIV. We consider it better to postpone publication of the Moriori results, until a fuller study has been made of all the craniological characters of the Royal College series, which we hope may soon be undertaken.

Tables A—D give the measurements made for the apes only. The tape measurements of the mesodacryal arc scarcely ever permitted of the determination to tenths of a millimetre, and this makes the tenths in the mesodacryal chord of small value, when we are determining the arcual index β. In the tables doubt as to a value is only expressed in the columns of the absolute measurements and not repeated in the index columns deduced from these absolute measurements. Of course the index values have no more weight than the data from which they are deduced.

Note. The Editor is responsible for the terminology used in this paper. He fully admits its barbarism. But no English form for the Greek adjective exists, and when a distinguished classical scholar fails to find one, science must take the matter into its own hands, and adjectives in -otic seem the only possibility. *Simotic* has practically Greek sanction; the pathologists have already adopted *stenotic*; thus the step to *mesotic* was not a very great one, and it is shorter than *mesomorphic* or *mesorhinal*,—were indeed compounds in rhinal not already allocated.

TABLE A. *Characters of the*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices				Simotic Chord	Simotic Subtense
					a	a'	$a-a'$	β		
1011. S	♂	19·5	29	9·1	46·7	49·4	− 2·7	48·7	6·0	3·3
7. 1. 8. 2	♂	21·6	30	9·9	45·8	43·1	+ 2·7	38·9	5·8	4·3
7. 1. 8. 1	♂	29·3	34	7·6	25·9	25·9	0	16·1	8·1	1·8
—	♂	33·5 ?	41 ?	7·6 ?	22·7	31·2	− 8·5	22·4	—	—
1011. u	♂	21·1	32	11·4	54·0	51·3	+ 2·7	51·7	5·7	2·7
1011. c	♂	16·2	23	7·5	46·3	45·1	+ 1·2	42·0	—	—
Okwa	♀	25·4	31·5	8·6	33·9	32·4	+ 1·5	24·0	—	—
7. 1. 8. 4	♀	19·1	29·2	9·2	48·2	52·1	− 3·9	52·9	3·6	2·2
7. 1. 8. 5	♀	22·2	31·7	9·6	43·2	45·6	− 2·4	42·8	5·4	3·3
7. 1. 8. 3	♀	25·6 ?	33 ?	9·0	35·2	36·1	− ·9	28·9	4·7	3·8
7. 1. 8. 7	♀	25·0	32	9·2	36·8	35·4	+ 1·4	28·0	5·6	2·6
7. 1. 8. 6	♀	21·8 ?	30·6 ?	10·3	47·2	44·1	+ 3·1	40·4	4·0 ?	1·8 ?
64. 12. 1. 1	♀	17·8 ?	28 ?	10·0	56·2	54·8	+ 1·4	57·3	—	—
1011. d	♀	21·3	32·5	10·2	47·9	51·9	− 4·0	52·6	4·2	2·6
—	♀	21·5 ?	34 ?	10·1 ?	47·0	55·3	− 8·3	58·1	7·5 ?	5·7 ?
1011. b	♀	16·2	23	7·5	46·3	45·1	+ 1·2	42·0	3·0	2·0
22	♂	16·8	27 ?	11·0	65·5	56·9	+ 8·6	60·7	4·1	3·6
20	♂	18·7	23	5·3	28·3	31·6	− 3·3	23·0	—	—
21	♂	25·9	37	11·4	44·0	45·7	− 1·7	42·9	8·8	3·7
28	♂	22·6	30	8·7	38·5	38·8	− ·3	32·7	5·0	2·4
25	♂	22·0	33	8·6	39·1	50·3	−11·2	50·0	3·8	2·7
21. a	♂	25·0	30	7·2	28·8	29·2	− ·4	20·0	8·2	3·6
23	♀	23·8 ?	34 ?	10·5 ?	44·1	45·7	− 1·6	42·9	6·4	2·0
26	♀	24·2	31	8·3	34·2	35·5	− 1·3	28·1	6·0	2·6
27	♀	12·5 ?	19 ?	6·2 ?	49·6	45·7	+ 3·9	52·0	—	—
24	♀	22·0	28	8·4	38·2	34·9	+ 3·3	27·3	9·1	2·2
93	♂	20·3	29	8·4	41·4	45·7	− 4·3	42·9	8·7	3·9
—	♂	26·5	32	8·9	33·6	29·9	+ 3·7	20·8	6·0	3·1
T	♂	23·2	30	9·2	39·7	36·4	+ 3·3	29·3	8·2	3·7
T	♀	21·7	32	10·7	49·3	48·7	+ ·6	47·5	7·0	1·5 ?
—	♀	21·2	34	11·2	52·8	56·7	− 3·9	60·4	5·4	2·7
3	♀	19·0	23	5·4	28·4	30·1	− 1·7	21·1	4·2	1·1
6. 124	♀	11·8	17	6·4	54·2	46·5	+ 7·7	44·1	5·6	3·3
4	♀	19·5	31	11·6	59·5	55·8	+ 3·7	59·0	5·6	3·5
43. 5	♀	24·8 ?	30 ?	8·1	32·7	30·0	+ 2·7	21·0	—	—
"Gerrard"	♀	21·7	26	7·6	35·0	29·1	+ 5·9	19·8	—	—
5 xii. *05*	♀	28·0 ?	29 ?	10·2	36·4	—	—	3·6	—	—
16	♀	16·2 ?	28 ?	9·9	61·1	64·1	− 3·0	72·8	3·2	3·2
2	♀	17·6 ?	25 ?	7·6	43·2	45·2	− 2·0	42·1	—	—
3	♂	25·4	34	9·4	37·0	39·6	− 2·6	33·9	9·2	3·8
1[a]	♂	22·0 ?	30 ?	9·8	44·6	41·4	+ 3·2	36·4	—	—
1161. F	♀	18·2	28	9·6	52·8	52·7	+ ·1	53·9	4·2	3·1
1161. A	♂	24·7	36·5	12·9	52·2	48·9	+ 3·3	47·8	3·2	3·4
423. 5	♂	17·0	23	7·2	42·4	40·6	+ 1·8	35·3	3·0	2·2
7	♂	23·8 ?	33 ?	10·0	42·0	42·9	− ·9	38·7	6·0	2·9
L	♂	30·0	36·5	8·5	28·3	30·6	− 2·3	21·7	8·8	2·9
8	♂	20·9	32·5	11·6	55·5	53·7	+ 1·8	55·5	4·0	4·1
10	♂	19·7	28	9·0	45·7	45·2	+ ·5	42·1	6·5	5·0
1	♂	21·0	30	8·9	42·4	45·7	− 3·3	42·9	8·0	3·5
1	♂	25·0 ?	32·5 ?	8·9	35·6	36·9	− 1·3	30·0	—	—
2052	♂	21·8 ?	28 ?	6·5	29·8	35·8	− 6·0	28·5	8·8	3·0
982	♂	21·0 ?	27 ?	9·7	46·2	35·9	+10·3	28·6	—	—
2051	♀	15·4	23	6·9	44·8	49·9	− 5·1	49·4	6·2	4·3

Nasal Bridge in the Gorilla.

Simotic Index	Locus	Remarks
55·0	Natural History Museum	Sutures very vague
74·1	" " "	From Okuni district. Sutures clear
22·2	" " "	From Okuni district. Sutures vague [specimen mounted in case
—	" " "	Measurements uncertain owing to bad light and position of
47·4	" " "	Very large and heavy skull. Sutures not clear
—	" " "	Permanent teeth not quite erupted. Nasal bones too narrow to
—	" " "	Nasal bones too narrow to measure [measure
47·4	" " "	From Okuni district. Sutures vague
61·1	" " "	From Okuni district. Sutures clear
80·9	" " "	From Okuni district. Sutures vague
46·4	" " "	From Okuni district. Sutures very vague
45·0	" " "	From Okuni district. Sutures vague
—	" " "	In the Du Chaillu collection. Sutures obliterated
61·9	" " "	Sutures clear [specimen mounted in case
76·0	" " "	Measurements uncertain owing to bad light and position of
66·7	" " "	Young. Permanent teeth not quite erupted. Nasal bones
87·8	Royal College of Surgeons	Sutures clear [very narrow
—	" " "	Sutures clear. Split skull and therefore measurements uncertain. Nasal bone abnormal on one side
42·1	" " "	Heavily made skull. Sutures clear. Nasal and maxillary bones large, lacrymal bones very small
48·0	" " "	Young. Permanent teeth erupting
71·1	" " "	Split skull
43·9	" " "	Sutures clear [sutures are vague
31·3	" " "	Difficult to measure as skull is attached to skeleton in case and
43·3	" " "	Sutures vague
—	" " "	Adolescent. Split skull with badly fitting halves. Sutures
24·2	" " "	Perfect sutures [complex
44·8	University College, Zoology Department	Young adult. Sutures clear
51·7	" " " " "	Sutures rather vague
45·1	Professor Thane	Sutures vague
21·4	" "	Sutures very vague
50·0	Professor Pearson, Biometric Laboratory	Sutures vague
26·2	Cambridge Anthropological Laboratory	Sutures clear
58·9	" " " " "	Young adult. Sutures clear
62·5	" " " " "	Sutures clear
—	" " " " "	Sutures obliterated. Skull varnished
—	" " " " "	Sutures obliterated
—	" " " " "	Young. Third molar and canines not yet erupted. Nose apparently damaged and deformed. Omitted from statistical
100·0	" " " " "	Sutures obliterated. A high, narrow nose [reductions
—	" " " " "	Sutures obliterated. Fourth molar appearing on right side
41·3	" " " " "	Sutures clear. Bicuspids out of place on right side
—	" " " " "	Sutures very vague
73·8	Cambridge Zoological Laboratory ...	Sutures clear [bones
106·3	" " " " ...	Sutures clear. Very narrow nasal bones and large maxillary
73·3	Dr W. L. H. Duckworth	Young. Third molars and canines erupting
48·3	" " "	Sutures obliterated
33·0	" " "	Sutures clear but complex. Heavily developed skull
102·5	" " "	Sutures clear, but measurements rather uncertain on account of ossicles in naso-maxillary suture
76·9	" " "	Sutures obliterated. Much developed crest
43·8	" " "	Sutures clear. Heavily developed skull
—	Oxford Museum	Sutures obliterated. Cleft in sagittal cres
34·1	" "	Sutures obliterated
—	" "	Aged. Sutures obliterated
69·4	" "	Nearly adult. Sutures clear. Third molar nearly erupted

TABLE A. *Characters of the Nasal Bridge in the Gorilla.*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices				Simotic Chord	Simotic Subtense	Simotic Index	Locus	Remarks
					α	α′	α − α′	β					
1011 S	♂	19·5	29	9·1	46·7	49·4	2·7	48·7	6·0	3·3	55·0	Natural History Museum	Sutures very vague
7. 1 S. 2	♂	21·6	30	9·9	45·8	43·1	+ 2·7	38·9	5·8	4·3	74·1	" " "	From Okuni district. Sutures clear
7. 1 S. 1	♂	29·3	34	7·6	25·9	25·9	0	16·1	8·1	1·8	22·2	" " "	From Okuni district. Sutures vague [specimen mounted in case
	♂	33·5?	41?	7·6?	22·7	31·2	− 8·5	22·4	—	—	·	" " "	Measurements uncertain owing to bad light and position of
1011 a	♂	21·1	32	11·4	54·0	51·3	+ 2·7	51·7	5·7	2·7	47·4	" " "	Very large and heavy skull. Sutures not clear
1011 c	♂	16·2	23	7·5	46·3	45·1	+ 1·2	42·0	—		—	" " "	Permanent teeth not quite erupted. Nasal bones too narrow to
Okwa	♀	25·4	31·5	8·6	33·9	32·4	+ 1·5	24·0	—	—	—	" " "	Nasal bones too narrow to measure [measure
1. S. 4	♀	19·1	29·2	9·2	48·2	52·1	− 3·9	52·9	3·6	2·2	47·4	" " "	From Okuni district. Sutures vague
1. S. 5	♀	22·2	31·7	9·6	43·2	45·6	2·4	42·8	5·4	3·3	61·1	" " "	From Okuni district. Sutures clear
1. S. 3	♀	25·6?	33?	9·0	35·2	36·1	− ·9	28·9	4·7	3·8	80·9	" " "	From Okuni district. Sutures vague
1. S. 7	♀	25·0	32	9·2	36·8	35·4	+ 1·4	28·0	5·6	2·6	46·4	" " "	From Okuni district. Sutures very vague
1. S. 6	♀	21·8?	30·6?	10·3	47·2	44·1	+ 3·1	40·4	4·0?	1·8?	45·0	" " "	From Okuni district. Sutures vague
64. 12. 1. 1	♀	17·8?	28?	10·0	56·2	54·8	+ 1·4	57·3	—	—	—	" " "	In the Du Chaillu collection. Sutures obliterated
1011 d	♀	21·3	32·5	10·2	47·9	51·9	− 4·0	52·6	4·2	2·6	61·9	" " "	Sutures clear [specimen mounted in case
	♀	21·5?	34?	10·1?	47·0	55·3	− 8·3	58·1	7·5?	5·7?	76·0	" " "	Measurements uncertain owing to bad light and position of
1011. b	♀	16·2	23	7·5	46·3	45·1	+ 1·2	42·0	3·0	2·0	66·7	" " "	Young. Permanent teeth not quite erupted. Nasal bones
22	♂	16·8	27?	11·0	65·5	56·9	+ 8·6	60·7	4·1	3·6	87·8	Royal College of Surgeons	Sutures clear [very narrow
20	♂	18·7	23	5·3	28·3	31·6	3·3	23·0	—	-		" " "	Sutures clear. Split skull and therefore measurements uncertain. Nasal bone abnormal on one side
21	♂	25·9	37	11·4	44·0	45·7	− 1·7	42·9	8·8	3·7	42·1	" " "	Heavily made skull. Sutures clear. Nasal and maxillary bones large, lacrymal bones very small
28	♂	22·6	30	8·7	38·5	38·8	− ·3	32·7	5·0	2·4	48·0	" " "	Young. Permanent teeth erupting
[illegible]	♂	22·0	33	8·6	39·1	50·3	−11·2	50·0	3·8	2·7	71·1	" " "	Split skull
21. a	♂	25·0	30	7·2	28·8	29·2	− ·4	20·0	8·2	3·6	43·9	" " "	Sutures clear [sutures are vague
[illegible]	♀	23·8?	34?	10·5?	41·1	45·7	− 1·6	42·9	6·4	2·0	31·3	" " "	Difficult to measure as skull is attached to skeleton in case and
[illegible]	♀	24·2	31	8·3	34·2	35·5	− 1·3	28·1	6·0	2·6	43·3	" " "	Sutures vague
[illegible]	♀	12·5?	19?	6·2?	19·6	45·7	+ 3·9	52·0	—	-		" " "	Adolescent. Split skull with badly fitting halves. Sutures
[illegible]	♀	22·0	28	8·1	38·2	34·9	+ 3·3	27·3	9·1	2·2	24·2	" " "	Perfect sutures [complex
[illegible]	♂	20·3	29	8·4	41·4	45·7	− 4·3	42·9	8·7	3·9	44·8	University College, Zoology Department	Young adult. Sutures clear
	♂	26·5	32	8·9	33·6	29·9	+ 3·7	20·8	6·0	3·1	51·7	" " " " "	Sutures rather vague
T	♂	23·2	30	9·2	39·7	36·4	+ 3·3	29·3	8·2	3·7	45·1	Professor Thane	Sutures vague
T	♀	21·7	32	10·7	49·3	48·7	+ ·6	47·5	7·0	1·5?	21·4	" "	Sutures very vague
·	♀	21·2	34	11·2	52·8	56·7	− 3·9	60·4	5·4	2·7	50·0	Professor Pearson, Biometric Laboratory	Sutures vague
·	♀	19·0	23	5·4	28·4	30·1	− 1·7	21·1	4·2	1·1	26·2	Cambridge Anthropological Laboratory	Sutures clear
6. 124	♀	11·8	17	6·4	54·2	46·5	+ 7·7	44·1	5·6	3·3	58·9	" " " " "	Young adult. Sutures clear
4	♀	19·5	31	11·6	59·5	55·8	+ 3·7	59·0	5·6	3·5	62·5	" " " " "	Sutures clear
43. 5	♀	24·8?	30?	8·1	32·7	30·0	+ 2·7	21·0	—	—	—	" " " " "	Sutures obliterated. Skull varnished
"Gerrard"	♀	21·7	26	7·6	35·0	29·1	+ 5·9	19·8	—	—	—	" " " " "	Sutures obliterated
5 xii. 05	♀	28·0?	29?	10·2	36·4	—	—	3·6	—	—	—	" " " " "	Young. Third molar and canines not yet erupted. Nose apparently damaged and deformed. Omitted from statistical
16	♀	16·2?	28?	9·9	61·1	64·1	− 3·0	72·8	3·2	3·2	100·0	" " " " "	Sutures obliterated. A high, narrow nose [reductions
2	♀	17·6?	25?	7·6	43·2	45·2	− 2·0	42·1		—		" " " " "	Sutures obliterated. Fourth molar appearing on right side
·	♂	25·4	31	9·4	37·0	39·6	2·6	33·9	9·2	3·8	41·3	" " " " "	Sutures clear. Bicuspids out of place on right side
18	♂	22·0?	30?	9·8	41·6	41·4	+ 3·2	36·4	—	—	—	" " " " "	Sutures very vague
1161 F	♀	18·2	28	9·6	52·8	52·7	+ ·1	53·9	4·2	3·1	73·8	Cambridge Zoological Laboratory ...	Sutures clear [bones
1161 A	♂	21·7	36·5	12·9	52·2	48·9	+ 3·3	47·8	3·2	3·4	106·3	" " " " .	Sutures clear. Very narrow nasal bones and large maxillary
425. 5	♂	17·0	23	7·2	42·4	40·6	+ 1·8	35·3	3·0	2·2	73·3	Dr W. L. H. Duckworth	Young. Third molars and canines erupting
[illegible]	♂	23·8?	33?	10·0	42·0	42·9	− ·9	38·7	6·0	2·9	48·3	" " "	Sutures obliterated
L	♂	30·0	36·5	8·5	28·3	30·6	− 2·3	21·7	8·8	2·9	33·0	" " "	Sutures clear but complex. Heavily developed skull
[illegible]	♂	20·9	32·5	11·6	55·5	53·7	+ 1·8	55·5	4·0	4·1	102·5	" " "	Sutures clear, but measurements rather uncertain on account of ossicles in naso-maxillary suture
10	♂	19·7	28	9·0	45·7	45·2	+ ·5	42·1	6·5	5·0	76·9	" " "	Sutures obliterated. Much developed crest
1	♂	21·0	30	8·9	42·4	45·7	− 3·3	42·9	8·0	3·5	43·8	" " "	Sutures clear. Heavily developed skull
1	♂	25·0?	32·5?	8·9	35·6	36·9	− 1·3	30·0	—	—		Oxford Museum	Sutures obliterated. Cleft in sagittal cres
20. 2	♂	21·8?	28?	6·5	29·8	35·8	− 6·0	28·5	8·8	3·0	34·1	" "	Sutures obliterated
98. 2	♂	21·0?	27?	9·7	46·2	35·9	+10·3	28·6	—	—	—	" "	Aged. Sutures obliterated
2051	♂	15·4	23	6·9	44·8	49·9	− 5·1	49·4	6·2	4·3	69·4	" "	Nearly adult. Sutures clear. Third molar nearly erupted

TABLE B. *Characters of the*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices				Simotic Chord	Simotic Subtense
					α	α'	$\alpha - \alpha'$	β		
7	♂	20·0	27	7·1	35·5	40·4	− 4·9	35·0	6·0	·8
10	♂	14·8	16·8	3·8	25·7	23·5	+ 2·2	13·5	8·6	2·5
11. A	♂	15·2	19	4·8	31·6	33·2	− 1·6	25·0	5·0	1·3
11. 2	♂	15·1	18·5	2·7	17·9	31·2	− 13·3	22·5	9·8	1·8
10. A	♂	17·2	21	4·6	26·7	30·9	− 4·2	22·1	—	—
8	♂	16·0	19	3·9	24·4	28·3	− 3·9	18·8	8·2	2·1
9	♂	19·9	23	5·5	27·6	25·5	+ 2·1	15·6	7·7	·6
12	♂	21·1	27	5·5	26·1	35·4	− 9·3	28·0	10·0	1·0
3	♂	19·0	24	5·5	29·0	34·2	− 5·2	26·3	8·0	1·8
1	♂	19·2 ?	22 ?	4·9 ?	25·5	24·6	+ ·9	14·6	—	—
2	♀	19·0 ?	20·2 ?	4·0	21·1	15·7	+ 5·4	6·3	—	—
11	♀	23·0	25	5·4	23·5	18·6	+ 4·9	8·7	8·7	2·4
11. 1	♀	19·0	22 ?	5·0	26·3	25·6	+ ·7	15·8	9·0	1·6
11. B	♀	12·0	15	3·8	31·7	33·2	− 1·5	25·0	7·0	1·0
4	♀	17·0	19·5	2·7	15·9	24·7	− 8·8	14·7	8·7	·6
94	?♂	22·3	26	5·2	23·3	26·4	− 3·1	16·6	—	—
7. 7. 8. 19	♂	25·0 ?	31	6·8	27·2	32·4	− 5·2	24·0	—	—
94. 7. 25. 2	♂	21·0	23·7	6·4	30·5	23·0	+ 7·5	12·9	8·0	3·0
64. 12. 1. 7	?♀	19·0	22	5·4	28·4	25·6	+ 2·8	15·8	9·0	·9
87. 12. 1. 1	♂	14·1	17	4·8	34·0	29·7	+ 4·3	20·6	7·4	1·5
2. e	?♂	20·0	27	7·6	38·0	40·4	− 2·4	35·0	8·2	1·7
82. 9. 18. 1	?♂	22·0	25 ?	3·7 ?	16·8	23·6	− 6·8	13·6	—	—
64. 12. 1. 6	?♂	18·2	22	6·3 ?	34·6	30·0	+ 4·6	20·9	8·4	2·0
Koo Loo *2. 1*	?♂	20·5	26·5 ?	7·8	38·1	36·4	+ 1·7	29·3	—	—
Talbot *33*	?♂	18·2	24 ?	6·1	33·5	38·2	− 4·7	31·9	—	—
90. 6. 8. 1	♂	19·2 ?	24·2 ?	6·5	33·9	33·9	·0	26·0	—	—
65/11 (b)	?♂	16·0	18	3·7	23·1	22·6	+ ·5	12·5	—	—
65/11 (a)	?♀	22·0	28·5	6·2	28·2	36·5	− 8·3	29·5	—	—
1. 8. 9. 84	♀	15·2	21	6·0	39·5	42·6	− 3·1	38·2	—	—
7. 1. 8. 8	?♀	18·0	21·5	5·4	30·0	28·7	+ 1·3	19·4	—	—
1. 8. 9. 10	♂	21·0	25	5·9	28·1	28·5	− ·4	19·1	—	—
94. 7. 25. 3	?♀	20·0	25·5	7·3	36·5	35·1	+ 1·4	27·5	7·0	2·0
95. 4. 11	♀	21·0	26	7·3	34·8	32·3	+ 2·5	23·8	—	—
2. a	♂	19·8	24	6·2	31·3	30·2	+ 1·1	21·2	—	—
83	♀	11·0	13	2·4	21·8	27·7	− 5·9	18·2	8·2	1·8
90. 6. 8. 2	♀	14·3	18	4·5 ?	31·5	33·9	− 2·4	25·9	7·0	1·9
61. 7. 29. 12	♂	23·2	28	4·9	21·1	29·8	− 8·7	20·7	—	—
83. 7. 28. 18	♀	18·8	25·5 ?	5·8	30·9	40·8	− 9·9	35·6	—	—
87. 12. 1. 2	♂	19·0	24·5	7·3	38·4	36·1	+ 2·3	28·9	—	—
A	♀	19·0	23	5·3	27·9	30·1	− 2·2	21·1	8·0	1·0
B	?♂	16·2	20	5·3	32·7	32·0	+ ·7	23·5	12·8	3·0
C	?♀	21·0 ?	24 ?	5·3	25·2	24·3	+ ·9	14·3	—	—
D	♂	8·8	13	3·5	39·8	48·8	− 9·0	47·7	—	—
K	♂	20·0 ?	25 ?	6·0	30·0	33·2	− 3·2	25·0	—	—
5	♀	13·2	19	5·1	38·6	46·4	− 7·8	43·9	—	—
E	♂	19·6 ?	23·5 ?	6·3 ?	32·1	29·2	+ 2·9	19·9	—	—
1163	?♂	22·8 ?	28·5 ?	7·0 ?	30·7	33·2	− 2·5	25·0	—	—
1	♂	19·5	21	3·9	20·0	17·4	+ 2·6	7·7	—	—
2	?♂	13·0	18	4·6	35·4	42·8	− 7·4	38·5	—	—
1	?♂	15·0	18	3·6	24·0	29·2	− 5·2	20·0	—	—
2049[b]	♀	14·8	17·4	4·7	31·8	27·2	+ 4·6	17·6	—	—
979	♂	19·1 ?	21·5 ?	5·2 ?	27·2	22·7	+ 4·5	12·6	—	—

Nasal Bridge in the Chimpanzee.

Simotic Index	Locus	Remarks
13·3	Royal College of Surgeons	Adolescent. Sutures clear
29·1	,, ,, ,, ...	Sutures clear. A wide nose with abnormal nasal bones
26·0	,, ,, ,, ...	Nearly adult. Sutures clear. Nasal bones narrow
18·4	,, ,, ,, ...	Two sutures clear but the third suture quite obliterated
—	,, ,, ,, ...	Sutures obliterated
25·6	,, ,, ,, ...	Sutures very vague
7·8	,, ,, ,, ...	Sutures very vague. Split skull
10·0	,, ,, ,, ...	Sutures fairly clear
22·5	,, ,, ,, ...	Sutures very vague
—	,, ,, ,, ...	Sutures quite obliterated. Measurements uncertain owing to the position of the skull, which is attached to the skeleton in an exhibition case
—	,, ,, ,, ...	Sutures obliterated. Measurements uncertain owing to the position of the skull, which is attached to the skeleton [in an exhibition case
27·6	,, ,, ,, ...	Aged
17·8	,, ,, ,, ...	Sutures very vague
14·3	,, ,, ,, ...	Young adult. Sutures clear
6·9	,, ,, ,, ...	Two of the sutures clear
—	University College, Zoology Department	Sutures clear
—	Natural History Museum	Sutures obliterated. A large skull with very broad nose
37·5	,, ,, ,, ...	Sutures clear
10·0	,, ,, ,, ...	Sutures clear
20·3	,, ,, ,, ...	Young. Three molars erupting. Sutures clear
20·7	,, ,, ,, ...	Sutures fairly clear
—	,, ,, ,, ...	Sutures obliterated
23·8	,, ,, ,, ...	Sutures obliterated
—	,, ,, ,, ...	Sutures obliterated
—	,, ,, ,, ...	Sutures rather vague
—	,, ,, ,, ...	Sutures obliterated
—	,, ,, ,, ...	Young. Third molar erupted
—	,, ,, ,, ...	Sutures obliterated
—	,, ,, ,, ...	Aged. Sutures obliterated
—	,, ,, ,, ...	Aged. Sutures obliterated
—	,, ,, ,, ...	Aged. Sutures obliterated
28·6	,, ,, ,, ...	Aged. Sutures rather vague
—	,, ,, ,, ...	Sutures obliterated
—	,, ,, ,, ...	Sutures vague
22·0	,, ,, ,, ...	Sutures clear
27·1	,, ,, ,, ...	Sutures vague
—	,, ,, ,, ...	Sutures obliterated
—	,, ,, ,, ...	Sutures obliterated
—	,, ,, ,, ...	Sutures obliterated
12·5	Cambridge Anthropological Laboratory	Young. Third molar erupting. Sutures clear
23·4	,, ,, ,, ,, ,,	Young. Third molar erupting. Very broad nasal bones and [narrow maxillary
—	,, ,, ,, ,, ,,	Sutures obliterated
—	,, ,, ,, ,, ,,	Young. Third molar erupting. Sutures obliterated. Omitted from statistical reductions owing to incomplete develop- [ment
—	,, ,, ,, ,, ,,	Aged. Sutures obliterated
—	,, ,, ,, ,, ,,	Young adult. Sutures clear
—	,, ,, ,, ,, ,,	Sutures obliterated
—	Cambridge Zoological Laboratory ...	Sutures obliterated
—	Dr Duckworth	Sutures clear
—	,, ,, ...	Young. Third molar and canines erupting
—	Oxford Museum	Fronto-maxillary suture obliterated
—	,, ,, ...	Sutures clear
—	,, ,, ...	Sutures obliterated

TABLE B. *Characters of the Nasal Bridge in the Chimpanzee.*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices				Simotic Chord	Simotic Subtense	Simotic Index	Locus	Remarks
					α	α′	α − α′	β					
[illegible]	♂	20·0	27	7·1	35·5	40·4	4·9	35·0	6·0	·8	13·3	Royal College of Surgeons	Adolescent. Sutures clear
10	♀	14·8	16·8	3·8	25·7	23·5	+ 2·2	13·5	8·6	2·5	29·1	,, ,, ,, ...	Sutures clear. A wide nose with abnormal nasal bones
11. A	♀	15·2	19	1·8	31·6	33·2	− 1·6	25·0	5·0	1·3	26·0	,, ,, ,, ...	Nearly adult. Sutures clear. Nasal bones narrow
11. [illegible]	♂	15·1	18·5	2·7	17·9	31·2	− 13·3	22·5	9·8	1·8	18·4	,, ,, ,, ...	Two sutures clear but the third suture quite obliterated
10. A	♂	17·2	21	4·6	26·7	30·9	− 4·2	22·1	—	—	—	,, ,, ,, ...	Sutures obliterated
8	♂	16·0	19	3·9	24·4	28·3	− 3·9	18·8	8·2	2·1	25·6	,, ,, ,, ...	Sutures very vague
9	♂	19·9	23	5·5	27·6	25·5	+ 2·1	15·6	7·7	·6	7·8	,, ,, ,, ...	Sutures very vague. Split skull
12	♂	21·1	27	5·5	26·1	35·4	− 9·3	28·0	10·0	1·0	10·0	,, ,, ,, ...	Sutures fairly clear
3	♂	19·0	24	5·5	29·0	34·2	− 5·2	26·3	8·0	1·8	22·5	,, ,, ,, ...	Sutures very vague
1	♂	19·2 ?	22 ?	4·9 ?	25·5	24·6	+ ·9	14·6	—	—	—	,, ,, ,, ...	Sutures quite obliterated. Measurements uncertain owing to the position of the skull, which is attached to the skeleton in an exhibition case
2	♀	19·0 ?	20·2 ?	4·0	21·1	15·7	+ 5·4	6·3	—	—	—	,, ,, ,, ...	Sutures obliterated. Measurements uncertain owing to the position of the skull, which is attached to the skeleton [in an exhibition case
11	♀	23·0	25	5·4	23·5	18·6	+ 4·9	8·7	8·7	2·4	27·6	,, ,, ,, ...	Aged
11 [illegible]	♀	19·0	22 ?	5·0	26·3	25·6	+ ·7	15·8	9·0	1·6	17·8	,, ,, ,, ...	Sutures very vague
11. B	♀	12·0	15	3·8	31·7	33·2	− 1·5	25·0	7·0	1·0	14·3	,, ,, ,, ...	Young adult. Sutures clear
4	♀	17·0	19·5	2·7	15·9	24·7	− 8·8	14·7	8·7	·6	6·9	,, ,, ,, ...	Two of the sutures clear
94	?♂	22·3	26	5·2	23·3	26·4	− 3·1	16·6	—	—	—	University College, Zoology Department	Sutures clear
[illegible] *8. 19*	♂	25·0 ?	31	6·8	27·2	32·4	− 5·2	24·0		—		Natural History Museum	Sutures obliterated. A large skull with very broad nose
94. [illegible]	♂	21·0	23·7	6·4	30·5	23·0	+ 7·5	12·9	8·0	3·0	37·5	,, ,, ,, ...	Sutures clear
64. 12. 1. [illegible]	?♀	19·0	22	5·4	28·4	25·6	+ 2·8	15·8	9·0	·9	10·0	,, ,, ,, ...	Sutures clear
87. 12. 1. 1	♂	14·1	17	4·8	34·0	29·7	+ 4·3	20·6	7·4	1·5	20·3	,, ,, ,, ...	Young. Three molars erupting. Sutures clear
[illegible] e	?♂	20·0	27	7·6	38·0	40·4	− 2·4	35·0	8·2	1·7	20·7	,, ,, ,, ...	Sutures fairly clear
82. 9. 18. 1	?♂	22·0	25 ?	3·7 ?	16·8	23·6	− 6·8	13·6	—	—	—	,, ,, ,, ...	Sutures obliterated
64. 12. 1. 6	?♂	18·2	22	6·3 ?	34·6	30·0	+ 4·6	20·9	8·4	2·0	23·8	,, ,, ,, ...	Sutures obliterated
Koo Loo [illegible]	?♂	20·5	26·5 ?	7·8	38·1	36·4	+ 1·7	29·3	—	—	—	,, ,, ,, ...	Sutures obliterated
Talbot [illegible]	?♂	18·2	24 ?	6·1	33·5	38·2	− 4·7	31·9	—	—	—	,, ,, ,, ...	Sutures rather vague
90. 6. 8. 1	♂	19·2 ?	24·2 ?	6·5	33·9	33·9	·0	26·0	—	—	—	,, ,, ,, ...	Sutures obliterated
65. 11. b	?♂	16·0	18	3·7	23·1	22·6	+ ·5	12·5	—	—	—	,, ,, ,, ...	Young. Third molar erupted
65. 11. a	?♀	22·0	28·5	6·2	28·2	36·5	− 8·3	29·5	—	—	—	,, ,, ,, ...	Sutures obliterated
1. 8. 9. 84	♀	15·2	21	6·0	39·5	42·6	− 3·1	38·2	—	—	—	,, ,, ,, ...	Aged. Sutures obliterated
[illegible] *1. 8. 8*	?♀	18·0	21·5	5·4	30·0	28·7	+ 1·3	19·4	—	—	—	,, ,, ,, ...	Aged. Sutures obliterated
1. 8. [illegible] *10*	♂	21·0	25	5·9	28·1	28·5	− ·4	19·1	—	—	—	,, ,, ,, ...	Aged. Sutures obliterated
94. [illegible]	?♀	20·0	25·5	7·3	36·5	35·1	+ 1·4	27·5	7·0	2·0	28·6	,, ,, ,, ...	Aged. Sutures rather vague
97. 4. 11	♀	21·0	26	7·3	34·8	32·3	+ 2·5	23·8	—	—	—	,, ,, ,, ...	Sutures obliterated
[illegible] a	♂	19·8	24	6·2	31·3	30·2	+ 1·1	21·2	—	—	—	,, ,, ,, ...	Sutures vague
[illegible]	♀	11·0	13	2·4	21·8	27·7	− 5·9	18·2	8·2	1·8	22·0	,, ,, ,, ...	Sutures clear
90. 6. 8. [illegible]	♀	14·3	18	4·5 ?	31·5	33·9	− 2·4	25·9	7·0	1·9	27·1	,, ,, ,, ...	Sutures vague
61. [illegible] *12*	♂	23·2	28	4·9	21·1	29·8	− 8·7	20·7	—	—	—	,, ,, ,, ...	Sutures obliterated
[illegible] *18*	♀	18·8	25·5 ?	5·8	30·9	40·8	− 9·9	35·6	—	—	—	,, ,, ,, ...	Sutures obliterated
87. 12. 1. [illegible]	♂	19·0	24·5	7·3	38·4	36·1	+ 2·3	28·9	—	—	—	,, ,, ,, ...	Sutures obliterated
A	♀	19·0	23	5·3	27·9	30·1	− 2·2	21·1	8·0	1·0	12·5	Cambridge Anthropological Laboratory	Young. Third molar erupting. Sutures clear
B	?♂	16·2	20	5·3	32·7	32·0	+ ·7	23·5	12·8	3·0	23·4	,, ,, ,, ,, ,,	Young. Third molar erupting. Very broad nasal bones and [narrow maxillary
C	?♀	21·0 ?	24 ?	5·3	25·2	24·3	+ ·9	11·3	—	—		,, ,, ,, ,, ,,	Sutures obliterated
D	♂	8·8	13	3·5	39·8	48·8	− 9·0	47·7	—	—	—	,, ,, ,, ,, ,,	Young. Third molar erupting. Sutures obliterated. Omitted from statistical reductions owing to incomplete develop- [ment
K	♂	20·0 ?	25 ?	6·0	30·0	33·2	− 3·2	25·0	—	—	—	,, ,, ,, ,, ,,	Aged. Sutures obliterated
5	♀	13·2	19	5·1	38·6	46·4	− 7·8	43·9	—	—	—	,, ,, ,, ,, ,,	Young adult. Sutures clear
E	♂	19·6 ?	23·5 ?	6·3 ?	32·1	29·2	+ 2·9	19·9	—	—	—	,, ,, ,, ,, ,,	Sutures obliterated
1169	?♂	22·8 ?	28·5 ?	7·0 ?	30·7	33·2	2·5	25·0	—	—		Cambridge Zoological Laboratory ...	Sutures obliterated
1	♂	19·5	21	3·9	20·0	17·4	+ 2·6	7·7	—	—	—	Dr Duckworth	Sutures clear
2	?♂	13·0	18	4·6	35·4	42·8	− 7·4	38·5	—	—	—	,, ,, ...	Young. Third molar and canines erupting
1	[illegible]	15·0	18	3·6	24·0	29·2	− 5·2	20·0	—	—		Oxford Museum	Fronto-maxillary suture obliterated
2049[b]	♀	14·8	17·4	4·7	31·8	27·2	+ 4·6	17·6	—	—	—	,, ,, ...	Sutures clear
979	♂	19·1 ?	21·5 ?	5·2 ?	27·2	22·7	+ 4·5	12·6	—	—	—	,, ,, ...	Sutures obliterated

TABLE C. *Characters of the*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices			
					a	a'	$a-a'$	β
39	♂	18·0?	36	11·8?	65·6	79·6	−14·0	100·0
41	♂	16·8	59?	12·6	75·0	160·1	−85·1	251·2
37	♂	15·2	34·2	10·7	70·4	93·4	−23·0	125·0
38	♂	13·8	37?	12·9?	93·5	116·6	−23·1	168·1
38. A	♂	17·5	35	8·6	49·1	79·6	−30·5	100·0
45	♂	11·8	33?	11·4	96·6	122·7	−26·1	179·7
42	♂	16·8	38?	11·0	65·5	94·1	−28·6	126·2
40. A	♀	10·0	25·3	7·7	77·0	108·5	−31·5	153·0
40. B	♀	13·4	35·5	9·1	67·9	114·9	−47·0	164·9
40. E	♀	11·6	23	6·8	58·6	78·7	−20·1	98·3
44	♀	13·8	27·3	9·2	66·7	78·4	−11·7	97·8
44. A	♀	12·7	27	9·7	76·4	86·6	−10·2	112·6
46	♀	12·0	28·2	9·3	77·5	98·9	−21·4	135·0
—	♂	16·6	29	8·5	51·2	65·2	−14·0	74·7
59. 8. 16. 2	♂	11·0	33	12·7	115·5	133·4	−17·9	200·0
1101	♂	12·0	38	14·7	122·5	142·2	−19·7	216·7
92. 11. 5. 14	♂	18·8	31	9·7	51·6	59·4	− 7·8	64·9
55	♂	17·1	26	7·7	45·0	51·6	− 6·6	52·1
1087	♀	11·4	31	10·9	95·6	118·6	−23·0	171·9
1179. a	♂	10·2	29	10·5	102·9	125·2	−22·3	184·3
3. b	♂	13·8	31?	10·0	72·5	93·2	−20·7	124·6
—	?♂	14·0	35?	14·5	103·6	106·9	− 3·3	150·0
3. i	♂	15·2	30·5	11·2	73·7	80·0	− 6·3	100·7
92. 11. 5. 4	♂	17·0	35	11·6	68·2	82·9	−14·7	105·9
3. b^2	?♂	14·0	33	12·4	88·6	99·3	−10·7	135·7
3. m	♀	13·4	24	10·5	78·4	67·8	+10·6	79·1
32	♂	17·7	40	13·1	74·0	94·0	−20·0	126·0
79. 11. 21. 24	♀	13·0?	33·5?	11·1	85·4	111·1	−25·7	157·7
52. 12. 26	♂	16·0?	35?	9·6	60·0	90·0	−30·0	118·8
8. 7. 17. 1	♂	14·8	26	9·5	64·2	65·8	− 1·6	75·7
98. 2. 28. 11	♂	15·0	32	10·7	71·3	87·0	−15·7	113·3
3. j	♂	19·7	34	10·6	53·8	64·0	−10·2	72·6
3. a	♂	18·0	37	13·4	74·4	82·8	− 8·4	105·6
3. d	♂	15·6?	39	11·9	76·3	106·9	−30·6	150·0
3. q^2	♂	12·4	25	9·4	75·8	80·5	− 4·7	101·6
3. c^2	♂	13·8	32?	10·9	79·0	97·2	−18·2	131·9
92. 11. 5. 5	♂	20·8	37	11·0	52·9	67·1	−14·2	77·9
68. 16. 42	♂	16·0	28·5?	10·2	63·8	67·2	− 3·4	78·1
92. 11. 5. 3	♂	20·8	32	8·4	40·4	52·7	−12·3	53·9
85. 20. 7. 1	♂	16·0	29?	9·0	56·3	69·0	−12·7	81·3
85. 2. a 7. 2	♀	10·8	25	9·5	88·0	97·0	− 9·0	131·5
4	♂	14·7	34?	11·8	80·3	96·9	−16·6	131·3
3	?♂	16·0	34	12·4	77·5	86·6	− 9·1	112·5
2	?♂	13·0	32·5?	11·4	87·7	106·9	−19·2	150·0
1	♂	15·7?	33?	12·2	77·7	85·3	− 7·6	110·2
5	?♀	10·5?	26?	9·7	92·4	105·6	−13·2	147·6
1160. A	♂	15·0?	26?	7·1	47·3	64·4	−17·1	73·3
1160. G	♂	13·0	32	11·7	90·0	104·9	−14·9	146·2
1160. D	♂	12·5	33	11·7	93·6	114·4	−20·8	164·0
1163. C	♀	12·0	26	8·4	70·0	88·9	−18·9	116·7
1136	?♀	9·8	34	12·0	122·4	157·9	−35·5	246·9
D. *1*	♂	15·0	28	9·3	62·0	72·1	−10·1	86·7
D. *2*	♂	15·0	32?	10·0	66·7	87·0	−20·3	113·3
2	♂	14·7?	32	10·3	70·1	89·4	−19·3	117·7
A	♂	17·5	43	15·7	89·7	104·6	−14·9	145·7
1^a	♀	11·0	27	8·4	76·4	104·5	−28·1	145·5
2043^a	♀	9·0	28?	10·1	112·2	139·2	−27·0	211·1
2043	♂	14·7	41?	15·3	104·1	122·3	−18·2	178·9
1	♂	15·7	35·5	11·9	75·8	94·0	−18·2	126·1

Nasal Bridge in the Orang-utan.

Locus	Remarks
Royal College of Surgeons	Split skull with ill-fitting halves
” ” ”	Sutures very vague
” ” ”	Sutures clear
” ” ”	Sutures clear
” ” ”	Sutures clear
” ” ”	Sutures clear
” ” ”	Sutures clear
” ” ”	Sutures vague. Teeth misplaced
” ” ”	Sutures clear
” ” ”	Young. Sutures clear. Third molars erupting
” ” ”	Adult, but one third molar never erupted. Sutures clear
” ” ”	Sutures clear
” ” ”	Sutures clear
University College, Zoology Department	Lacrymo-maxillary suture not very clear
Natural History Museum	Sutures clear [side
” ” ”	No nasal bones. Maxillary bones meeting with ossicle on right
” ” ”	Sutures fairly clear. Unusually broad nose for an orang
” ” ”	Very aged. A large skull with broad nose and much developed crests. Sutures vague
” ” ”	Fragment only. Sutures clear
” ” ”	In the Whale Room
” ” ”	
” ” ”	In the Whale Room
” ” ”	Sutures clear
” ” ”	Sutures clear
” ” ”	Aged. Sutures rather vague
” ” ”	Lacrymo-maxillary suture obliterated
” ” ”	Lacrymo-maxillary suture obliterated
” ” ”	Sutures clear. No nasal bones
” ” ”	Aged. Sutures clear [much developed crest
” ” ”	Aged. Sutures clear. Very long "spindle" nasal bone and
” ” ”	A large skull with rather vague sutures
” ” ”	Lacrymo-maxillary suture obliterated
” ” ”	Sutures clear
” ” ”	Sutures clear
” ” ”	Sutures clear. Nasal bones very narrow
” ” ”	Lacrymo-maxillary suture obliterated
” ” ”	Sutures clear. Very wide nose [bones extremely narrow
” ” ”	Broken orbits and therefore measurements uncertain. Nasal
” ” ”	Sutures clear. Very large skull
” ” ”	Sutures clear
” ” ”	Young adult. Sutures clear
Cambridge Anthropological Laboratory	Young adult. Sutures clear
” ” ” ”	Sutures clear
” ” ” ”	Young adult. Sutures clear
” ” ” ”	Lacrymo-maxillary suture absent, other sutures clear
” ” ” ”	Young. Third molar and canine erupting. Split skull and therefore uncertain measurements. Believed to be large species. Omitted from statistical reductions
Cambridge Zoological Laboratory ...	Right lacrymo-maxillary suture obliterated, left placed high on
” ” ” ...	Sutures clear [side of nose
” ” ” ...	Young adult. Four upper molars on both sides and four lower
” ” ” ...	Sutures clear [molars on left side
” ” ” ...	Young. Third molar and canines erupting. Believed to be large species. Omitted from statistical reductions
Dr Duckworth	Sutures clear
” ”	Sutures clear
” ”	Sutures clear
” ”	No nasal bones. Maxillary bones much developed
” ”	Sutures clear [incisors and canines
Oxford Museum	Young adult. Bicuspids on each side misplaced behind lateral
” ”	Young adult. Sutures clear
” ”	Sutures clear. ?Ossicle at dacryon on left side

TABLE C. *Characters of the*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices α	α′	α − α′	β
39	♂	18·0?	36	11·8?	65·6	79·6	−14·0	100·0
41	♂	16·8	59?	12·6	75·0	160·1	−85·1	251·2
37	♂	15·2	34·2	10·7	70·4	93·4	−23·0	125·0
38	♂	13·8	37?	12·9?	93·5	116·6	−23·1	168·1
38. A	♂	17·5	35	8·6	49·1	79·6	−30·5	100·0
45	♂	11·8	33?	11·4	96·6	122·7	−26·1	179·7
42	♂	16·8	38?	11·0	65·5	94·1	−28·6	126·2
40. A	♀	10·0	25·3	7·7	77·0	108·5	−31·5	153·0
40. B	♀	13·4	35·5	9·1	67·9	114·9	−47·0	164·9
40. E	♀	11·6	23	6·8	58·6	78·7	−20·1	98·3
44	♀	13·8	27·3	9·2	66·7	78·4	−11·7	97·8
44. A	♀	12·7	27	9·7	76·4	86·6	−10·2	112·6
46	♀	12·0	28·2	9·3	77·5	98·9	21·4	135·0
	♂	16·6	29	8·5	51·2	65·2	−14·0	74·7
59. 8. 16. 2	♂	11·0	33	12·7	115·5	133·4	−17·9	200·0
1101	♂	12·0	38	14·7	122·5	142·2	−19·7	216·7
92. 11. 5. 14	♂	18·8	31	9·7	51·6	59·4	−7·8	64·9
55	♂	17·1	26	7·7	45·0	51·6	−6·6	52·1
1087	♀	11·4	31	10·9	95·6	118·6	−23·0	171·9
1179. a	♂	10·2	29	10·5	102·9	125·2	22·3	184·3
3. b	♂	13·8	31?	10·0	72·5	93·2	−20·7	124·6
—	?♂	14·0	35?	14·5	103·6	106·9	−3·3	150·0
3. i	♂	15·2	30·5	11·2	73·7	80·0	−6·3	100·7
92. 11. 5. 4	♂	17·0	35	11·6	68·2	82·9	14·7	105·9
3. b2	?♂	14·0	33	12·4	88·6	99·3	−10·7	135·7
3. m	♀	13·4	24	10·5	78·4	67·8	+10·6	79·1
32	♂	17·7	40	13·1	74·0	94·0	−20·0	126·0
79. 11. 21. 24	♀	13·0?	33·5?	11·1	85·4	111·1	−25·7	157·7
52. 12. 26	♂	16·0?	35?	9·6	60·0	90·0	−30·0	118·8
8. 7. 17. 1	♂	14·8	26	9·5	64·2	65·8	1·6	75·7
98. 2. 28. 11	♂	15·0	32	10·7	71·3	87·0	−15·7	113·3
3. j	♂	19·7	34	10·6	53·8	64·0	−10·2	72·6
3. a	♂	18·0	37	13·4	74·4	82·8	−8·4	105·6
3. d	♂	15·6?	39	11·9	76·3	106·9	−30·6	150·0
3. q2	♂	12·4	25	9·4	75·8	80·5	−4·7	101·6
3. c2	♂	13·8	32?	10·9	79·0	97·2	−18·2	131·9
92. 11. 5. [illegible]	♂	20·8	37	11·0	52·9	67·1	14·2	77·9
68. 16. 42	♂	16·0	28·5?	10·2	63·8	67·2	−3·4	78·1
92. 11. 5. 3	♂	20·8	32	8·4	40·4	52·7	12·3	53·9
85. 20. 7. 1	♂	16·0	29?	9·0	56·3	69·0	−12·7	81·3
85. 2. [illegible]. 7. 2	♀	10·8	25	9·5	88·0	97·0	−9·0	131·5
4	♂	14·7	31?	11·8	80·3	96·9	−16·6	131·3
3	?♂	16·0	34	12·4	77·5	86·6	−9·1	112·5
2	?♂	13·0	32·5?	11·4	87·7	106·9	−19·2	150·0
1	♂	15·7?	33?	12·2	77·7	85·3	−7·6	110·2
6	?♀	10·5?	26?	9·7	92·4	105·6	−13·2	147·6
1160 A	♂	15·0?	26?	7·1	47·3	64·4	−17·1	73·3
1160 G	♂	13·0	32	11·7	90·0	104·9	−14·9	146·2
1160 D	♂	12·5	33	11·7	93·6	114·4	−20·8	164·0
1163 C	♀	12·0	26	8·4	70·0	88·9	−18·9	116·7
1156	?♀	9·8	34	12·0	122·4	157·9	−35·5	246·9
D. 1	♂	15·0	28	9·3	62·0	72·1	−10·1	86·7
D. 2	♂	15·0	32?	10·0	66·7	87·0	−20·3	113·3
2	♂	14·7?	32	10·3	70·1	89·4	−19·3	117·7
A	♂	17·5	43	15·7	89·7	104·6	−14·9	145·7
1a	♂	11·0	27	8·4	76·4	104·5	−28·1	145·5
204a	♀	9·0	28?	10·1	112·2	139·2	−27·0	211·1
204b	♂	14·7	41?	15·3	104·1	122·3	−18·2	178·9
1	♂	15·7	35·5	11·9	75·8	94·0	−18·2	126·1

Nasal Bridge in the Orang-utan.

Locus	Remarks
Royal College of Surgeons	Split skull with ill-fitting halves
„ „ „	Sutures very vague
„ „ „	Sutures clear
„ „ „	Sutures clear
„ „ „	Sutures clear
„ „ „	Sutures clear
„ „ „	Sutures vague. Teeth misplaced
„ „ „	Sutures clear
„ „ „	Young. Sutures clear. Third molars erupting
„ „ „	Adult, but one third molar never erupted. Sutures clear
„ „ „	Sutures clear
„ „ „	Sutures clear
University College, Zoology Department	Lacrymo-maxillary suture not very clear
Natural History Museum	Sutures clear [side
„ „ „	No nasal bones. Maxillary bones meeting with ossicle on right
„ „ „	Sutures fairly clear. Unusually broad nose for an orang
„ „ „	Very aged. A large skull with broad nose and much developed crests. Sutures vague
„ „ „	Fragment only. Sutures clear
„ „ „	In the Whale Room
„ „ „	
„ „ „	In the Whale Room
„ „ „	Sutures clear
„ „ „	Sutures clear
„ „ „	Aged. Sutures rather vague
„ „ „	Lacrymo-maxillary suture obliterated
„ „ „	Lacrymo-maxillary suture obliterated
„ „ „	Sutures clear. No nasal bones
„ „ „	Aged. Sutures clear [much developed crest
„ „ „	Aged. Sutures clear. Very long "spindle" nasal bone and
„ „ „	A large skull with rather vague sutures
„ „ „	Lacrymo-maxillary suture obliterated
„ „ „	Sutures clear
„ „ „	Sutures clear
„ „ „	Sutures clear. Nasal bones very narrow
„ „ „	Lacrymo-maxillary suture obliterated
„ „ „	Sutures clear. Very wide nose [bones extremely narrow
„ „ „	Broken orbits and therefore measurements uncertain. Nasal
„ „ „	Sutures clear. Very large skull
„ „ „	Sutures clear
„ „ „	Young adult. Sutures clear
Cambridge Anthropological Laboratory	Young adult. Sutures clear
„ „ „ „	Sutures clear
„ „ „ „	Young adult. Sutures clear
„ „ „ „	Lacrymo-maxillary suture absent, other sutures clear
„ „ „ „	Young. Third molar and canine erupting. Split skull and therefore uncertain measurements. Believed to be large species. Omitted from statistical reductions
Cambridge Zoological Laboratory ...	Right lacrymo-maxillary suture obliterated, left placed high on
„ „ „ ...	Sutures clear [side of nose
„ „ „ ...	Young adult. Four upper molars on both sides and four lower
„ „ „ ...	Sutures clear [molars on left side
„ „ „ ...	Young. Third molar and canines erupting. Believed to be large species. Omitted from statistical reductions
Dr Duckworth	Sutures clear
„ „	Sutures clear
„ „	Sutures clear
„ „	No nasal bones. Maxillary bones much developed
„ „	Sutures clear [incisors and canines
Oxford Museum	Young adult. Bicuspids on each side misplaced behind lateral
„ „	Young adult. Sutures clear
„ „	Sutures clear. ?Ossicle at dacryon on left side

TABLE D. *Characters of the Nasal Bridge*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices				Simotic Chord	Simotic Subtense
					α	α'	$\alpha-\alpha'$	β		
6. 10. 4. 1	♂	13·2	20	4·3	32·6	51·2	−18·6	51·5	—	—
38. 3. 13. 7	⚲	11·0	16	3·5	31·8	47·4	−15·6	45·5	5·5	·9
38. 3. 13. 5	⚲	9·0	10	1·2	13·3	21·2	− 7·9	11·1	4·3	·3
81. 3. 15. 1	♂	11·8	17	3·9	33·1	46·5	−13·4	44·1	6·6	·8
4 b	?♂	15·0	20	4·3	28·7	39·2	−10·5	33·3	5·4	·2
67. 11. 12. 5	⚲	15·8	22	3·9	24·7	43·3	−18·6	39·2	—	—
8. 7. 201	♂	12·5	17·5	3·5	28·0	43·8	−15·8	40·0	—	—
58	♂	16·0 ?	23·8 ?	5·9 ?	36·9	49·5	−12·6	48·8	8·0 ?	·7 ?
59	♂	14·2	20·5	4·5	31·7	46·7	−15·0	44·4	6·2	·8
64	♀	10·0	15	3·3	33·0	50·3	−17·3	50·0	5·6	·5
65	⚲	9·8	15	4·1	41·8	52·2	−10·4	53·1	5·5	·3
66	⚲	9·9	15	4·3	43·4	51·2	− 7·8	51·5	4·3	·4
70	⚲	11·0	16·5 ?	3·6	32·7	50·3	−17·6	50·0	5·8 ?	·3
71	♂	9·8	13·5	2·8	28·6	42·3	−13·7	37·8	6·0	·4
72	♀	9·0	14·2	3·6	40·0	55·1	−15·1	57·8	4·0	·5
(5. a) *43. 12. 7. 5*	⚲	11·4	16·5	4·2	36·8	46·9	−10·1	44·8	4·0 ?	·5
11. 2. 24. 4	♂	9·9	20	5·7	57·6	80·8	−23·2	102·0	6·0	1·0
8. 11. 1. 1	⚲	13·0 ?	22 ?	5·4	41·5	62·0	−20·5	69·2	—	—
90. 1. 28. 1	⚲	12·0	15·5	1·7	14·2	36·3	−22·1	29·2	6·0	− ·1
1097 a	⚲	11·8	17·3	3·9	33·1	48·1	−15·0	46·6	5·3	·1
1097 c	⚲	10·4	17·4 ?	4·6 ?	44·2	60·8	−16·6	67·3	—	—
9. 1. 51	♂	10·3 ?	17·3 ?	4·3	41·7	61·3	−19·6	68·0	4·3 ?	·2
1006 a	⚲	9·3	17	4·0	43·0	69·9	−26·9	82·8	5·7 ?	·1
'0. 10. 1. 8	♀	11·0 ?	18 ?	4·4	40·0	58·6	−18·6	63·6	—	—
10. 10. 1. 6	♂	10·7	18	3·1	29·0	61·4	−32·4	68·2	5·0 ?	·2 ?
79. 8. 30. 1	⚲	11·3	16	3·5	31·0	44·8	−13·8	41·6	6·0	·8
10. 10. 1. 7	♀	11·4	16·6	3·0	26·3	47·5	−21·2	45·6	5·2	·5
54. 10. 5. 1	⚲	12·8	20	3·7	28·9	54·2	−25·3	56·3	—	—
1006 d / *48. 1. 27. 46*	⚲	11·0 ?	19	4·5	40·9	64·0	−23·1	72·7	—	—
60. 5. 4. 48	⚲	9·0 ?	16	3·7	41·1	67·0	−25·9	77·8	—	—
60. 5. 4. 49	⚲	9·8	17	4·2	42·9	64·5	−22·6	73·5	—	—
1095 a	⚲	9·0	12	2·4	26·7	39·2	−12·5	33·3	5·9	·9
—	⚲	11·0	17·5	3·7	33·6	55·9	−22·3	59·1	—	—
1095 c	⚲	9·0	15·3	4·5	50·0	62·4	−12·4	70·0	6·6	·7

The symbol ⚲ denotes sex unknown.

in the Gibbons, Hylobates and Symphalangus.

Simotic Index	Locus	Remarks
—	Natural History Museum	Hylobates syndactylus. From Pahang. Sutures obliterated
16·4	” ” ”	From Sumatra. Very young
7·0	” ” ”	Symphalangus. From Sumatra. Young adult
12·1	” ” ”	Symphalangus. From Sumatra. Asymmetric lacrymal bones
3·7	” ” ”	Symphalangus syndactylus. From Sumatra. Old. Sutures vague
—	” ” ”	Symphalangus syndactylus. Sutures vague. Nasals absent at the bridge
—	” ” ”	Symphalangus syndactylus. From Pahang. Very old. Sutures quite obliterated
8·8	Royal College of Surgeons	Symphalangus. Old. Sutures obliterated
12·9	” ” ”	Symphalangus. Old. Sutures rather vague
8·9	” ” ”	Hylobates sp.
5·5	” ” ”	Hylobates sp. Young
9·3	” ” ”	Hylobates lar. Adult
5·2	” ” ”	Hylobates sp. Old. Sutures vague, nasal sutures obliterated
6·7	” ” ”	Hylobates sp. Young
12·5	” ” ”	Hylobates sp. Very young
12·5	Natural History Museum	Hylobates leuciscus from Malacca. Very old. Sutures obliterated
16·7	” ” ”	Hylobates hainanus from Hainan. Young. Peculiar lacrymal bones
—	” ” ”	Hylobates gabriella from Annam. Old. Sutures quite obliterated
−1·7	” ” ”	Hylobates mulleri from N. Borneo. Nasal sutures vague
1·9	” ” ”	Hylobates agilis
—	” ” ”	Hylobates agilis from Sumatra. Sutures obliterated
4·7	” ” ”	Hylobates leuciscus from Java. Sutures obliterated
1·8	” ” ”	Hylobates hoolock. Young. Apparently no nasals, but process of frontal
—	” ” ”	Hylobates lar. From Malay Peninsula. Sutures obliterated
4·0	” ” ”	Hylobates lar. From Malay Peninsula. Sutures obliterated
13·3	” ” ”	Hylobates leuciscus from Borneo
9·6	” ” ”	Hylobates lar. From Malay Peninsula
—	” ” ”	Hylobates lar. From Malacca. Sutures obliterated
—	” ” ”	Hylobates sp. Ridge on maxilla. Nasal sutures obliterated
—	” ” ”	Hylobates lar. From Malacca. Sutures obliterated
—	” ” ”	Hylobates lar. From Malacca. Nasals broken
15·3	” ” ”	Hylobates lar. Young
—	” ” ”	Hylobates lar. Sutures obliterated, bisected skull
10·6	” ” ”	Hylobates sp. Very young

TABLE D. *Characters of the Nasal Bridge in the Gibbons, Hylobates and Symphalangus.*

Museum Number	Sex	Meso-dacryal Chord	Meso-dacryal Arc	Meso-dacryal Subtense	Mesodacryal Indices α	α'	$\alpha - \alpha'$	β	Simotic Chord	Simotic Subtense
6. 10. 4. 1	♂	13·2	20	4·3	32·6	51·2	−18·6	51·5		
88. 3. 13. ?	⚲	11·0	16	3·5	31·8	47·4	15·6	45·5	5·5	·9
?8. 3. 15. ?	♂	9·0	10	1·2	13·3	21·2	− 7·9	11·1	4·3	·3
81. ?. 15. 1	♂	11·8	17	3·9	33·1	46·5	−13·4	44·1	6·6	·8
4 b	?⚲	15·0	20	4·3	28·7	39·2	−10·5	33·3	5·4	·2
?. 11. 12. ?	⚲	15·8	22	3·9	24·7	43·3	−18·6	39·2		
8. 7. 201	♂	12·5	17·5	3·5	28·0	43·8	−15·8	40·0	—	
58	♂	16·0 ?	23·8 ?	5·9 ?	36·9	49·5	12·6	48·8	8·0 ?	·7 ?
59	♂	14·2	20·5	4·5	31·7	46·7	−15·0	44·4	6·2	·8
64	♀	10·0	15	3·3	33·0	50·3	−17·3	50·0	5·6	·5
65	♂	9·8	15	4·1	41·8	52·2	10·4	53·1	5·5	·3
66	♂	9·9	15	4·3	43·4	51·2	− 7·8	51·5	4·3	·4
70	♂	11·0	16·5 ?	3·6	32·7	50·3	−17·6	50·0	5·8 ?	·3
71	♂	9·8	13·5	2·8	28·6	42·3	−13·7	37·8	6·0	·4
72	♀	9·0	14·2	3·6	40·0	55·1	15·1	57·8	4·0	·5
?. a 43. 12. ?. 5	♂	11·4	16·5	4·2	36·8	46·9	−10·1	44·8	4·0 ?	·5
11. 2. 24. 4	♂	9·9	20	5·7	57·6	80·8	−23·2	102·0	6·0	1·0
8. 11. 1. 1	♂	13·0 ?	22 ?	5·4	41·5	62·0	−20·5	69·2	·	
90. 1. 28. 1	♂	12·0	15·5	1·7	14·2	36·3	−22·1	29·2	6·0	−·1
1097 a	♂	11·8	17·3	3·9	33·1	48·1	−15·0	46·6	5·3	·1
1097 c	♂	10·4	17·4 ?	4·6 ?	44·2	60·8	−16·6	67·2		—
9. 1. 51	♂	10·3 ?	17·3 ?	4·3	41·7	61·3	−19·6	68·0	4·3 ?	·2
1006 a	♂	9·3	17	4·0	43·0	69·9	−26·9	82·8	5·7 ?	·1
'0. 10. 1. 8	♀	11·0 ?	18 ?	4·4	40·0	58·6	−18·6	63·6	—	
10. 10. 1. 6	♂	10·7	18	3·1	29·0	61·4	32·4	68·2	5·0 ?	·2 ?
79. 8. 30. 1	♂	11·3	16	3·5	31·0	44·8	13·8	44·6	6·0	·8
10. 10. 1. 7	♀	11·4	16·6	3·0	26·3	47·5	−21·2	45·6	5·2	·5
54. 10. 5. 1	♂	12·8	20	3·7	28·9	54·2	−25·3	56·3		·
1006 d / *48. 1. 27. 46*	♂	11·0 ?	19	4·5	40·9	64·0	−23·1	72·7		-
60. 5. 4. 48	⚲	9·0 ?	16	3·7	41·1	67·0	−25·9	77·8	—	—
60. 5. 4. 49	⚲	9·8	17	4·2	42·9	64·5	−22·6	73·5		-
1095 a	⚲	9·0	12	2·4	26·7	39·2	−12·5	33·3	5·9	·9
	⚲	11·0	17·5	3·7	33·6	55·9	−22·3	59·1	—	—
1095 c	⚲	9·0	15·3	4·5	50·0	62·4	12·4	70·0	6·6	·7

Simotic Index	Locus	Remarks
—	Natural History Museum	Hylobates syndactylus. From Pahang. Sutures obliterated
16·4	„ „ „	From Sumatra. Very young
7·0	„ „ „	Symphalangus. From Sumatra. Young adult
12·1	„ „ „	Symphalangus. From Sumatra. Asymmetric lacrymal bones
3·7	„ „ „	Symphalangus syndactylus. From Sumatra. Old. Sutures vague
	„ „ „	Symphalangus syndactylus. Sutures vague. Nasals absent at the bridge
—	„ „ „	Symphalangus syndactylus. From Pahang. Very old. Sutures quite obliterated
8·8	Royal College of Surgeons	Symphalangus. Old. Sutures obliterated
12·9	„ „ „	Symphalangus. Old. Sutures rather vague
8·9	„ „ „	Hylobates sp.
5·5	„ „ „	Hylobates sp. Young
9·3	„ „ „	Hylobates lar. Adult
5·2	„ „ „	Hylobates sp. Old. Sutures vague, nasal sutures obliterated
6·7	„ „ „	Hylobates sp. Young
12·5	„ „ „	Hylobates sp. Very young
12·5	Natural History Museum	Hylobates leuciscus from Malacca. Very old. Sutures obliterated
16·7	„ „ „	Hylobates hainanus from Hainan. Young. Peculiar lacrymal bones
	„ „ „	Hylobates gabriella from Annam. Old. Sutures quite obliterated
−1·7	„ „ „	Hylobates mulleri from N. Borneo. Nasal sutures vague
1·9	„ „ „	Hylobates agilis
—	„ „ „	Hylobates agilis from Sumatra. Sutures obliterated
4·7	„ „ „	Hylobates leuciscus from Java. Sutures obliterated
1·8	„ „ „	Hylobates hoolock. Young. Apparently no nasals, but process of frontal
	„ „ „	Hylobates lar. From Malay Peninsula. Sutures obliterated
4·0	„ „ „	Hylobates lar. From Malay Peninsula. Sutures obliterated
13·3	„ „ „	Hylobates leuciscus from Borneo
9·6	„ „ „	Hylobates lar. From Malay Peninsula
	„ „ „	Hylobates lar. From Malacca. Sutures obliterated
	„ „ „	Hylobates sp. Ridge on maxilla. Nasal sutures obliterated
	„ „ „	Hylobates lar. From Malacca. Sutures obliterated
—	„ „ „	Hylobates lar. From Malacca. Nasals broken
15·3	„ „ „	Hylobates lar. Young
-	„ „ „	Hylobates lar. Sutures obliterated, bisected skull
10·6	„ „ „	Hylobates sp. Very young

The symbol ⚲ denotes sex unknown.

ON THE CALCULATION OF INTRA-CLASS AND INTER-CLASS COEFFICIENTS OF CORRELATION FROM CLASS MOMENTS WHEN THE NUMBER OF POSSIBLE COMBINATIONS IS LARGE.

By J. ARTHUR HARRIS, Ph.D., Carnegie Institution of Washington, U.S.A.

I. Introductory Remarks.

If one designates as classes the groups of measurements or evaluations segregated out of the whole series under consideration, then those correlations which measure the similarity in character of the individuals of the same class may be designated as intra-class coefficients of correlation, while those which measure the resemblance between the individuals of different but (in some manner) related classes may be designated as inter-class coefficients of correlation.

Intra-class and inter-class should not be confused with intra-racial and inter-racial. These latter terms seem to have been first used in a memoir on Naquada Crania by C. D. Fawcett* and discussed for the local races of the lesser celandine† and in greater detail with anthropometric illustrations by Jacob, Lee, and Pearson‡. Examples of the practical use of inter-racial correlation coefficients are also given by Tschepourkowsky§. By intra-racial correlation as the term is used in these discussions, one understands those coefficients expressing the interdependence between the measures on a series of individuals of the same race. Under inter-racial are included the correlations between the types (means, modes) of a series of races.

In intra-class and inter-class correlation, on the other hand, the unit is always the individual (in the statistical sense), not the type of a series of individuals. The purpose of the correlation is to determine the resemblance for any series of characters of the individuals of sub-classes, or of an associated pair of sub-classes,

* *Biometrika*, Vol. I. pp. 460—461, 1902.

† *Biometrika*, Vol. II. pp. 152—153, 1903.

‡ *Biometrika*, Vol. II. pp. 347—356, 1903.

§ *Biometrika*, Vol. IV. pp. 161—168, 286—312, 1905.

as compared with random pairs from the population (or universe) which they constitute. Thus fraternal correlations are intra-class and also intra-racial, as are also homotypic coefficients. In one case, the "class" is fixed by ancestry and family environment, in the other, by ancestry and individual environment. Avuncular correlations—in which all nephews (nieces) of the descendant generation are compared with all the uncles (aunts) of the ascendant generation—are illustrations of inter-class, but also intra-racial, correlation.

The most familiar illustration of intra-class correlations are the fraternal and homotypic; practically the only inter-class relationships hitherto considered are the avuncular and the cousinship in heredity. But these correlations might be powerful research tools in morphology and physiology as well as in heredity and sociology.

The reason for their limited use in the past is in part technical. When the number of observations is as large as is desirable, the formation and verification of the ordinary correlation table (where each individual of the N used is entered only once) is an irksome task, but when each measurement is compared with a number of others, the purely clerical labour involved in tabulation becomes onerous in the extreme. If n be the number of individuals in any class, the number of combinations within the class is $\frac{1}{2}n(n-1)$, while if the tables be rendered symmetrical* as is generally desirable†, the number of combinations for each class is $n(n-1)$, which gives for the m classes constituting the population $S[n(n-1)]$ or, where n is constant, $m[n(n-1)]$ entries in the intra-class correlation surface. With n as low as 20 and m only 250, this gives 95,000 combinations. In inter-class correlation if p be the number in the first and q the number of observations in the second (associated) class, the permutations for a class are pq, and the entries for the population $S(pq)$, or mpq if both p and q are invariable. Obviously, neither m, p, nor q need be large to give very heavy tables.

The labour involved in constructing such tables and making sure that they are free from errors is very great indeed. To meet this difficulty in extreme cases, Pearson‡ has shown how formulae may be deduced for obtaining the coefficient of correlation between any grades of kindred from the means of arrays into which the kindred may be grouped. The method suggested here reduces the work to still simpler terms, obviating entirely the necessity for the

* For discussions of the advantages of the symmetrical table see K. Pearson and others, *Phil. Trans. Roy. Soc. Lond.* A, Vol. CXCVII. pp. 285—379, 1901; R. Pearl, *Biometrika*, Vol. v. pp. 249—251, 1907; K. Pearson and A. Barrington, *Eugenics Laboratory Memoirs*, No. V., 1909; J. Arthur Harris, *Biometrika*, Vol. VII. pp. 214—218, 1909.

† Jennings (*American Naturalist*, Vol. XLV. pp. 123—128, 1911) has suggested a method of calculating the symmetrical table constant without actually rendering the table symmetrical. I believe his method is not serviceable where more than a single pair of characters are involved.

‡ Pearson, K., *Phil. Trans. Roy. Soc. Lond.* A, Vol. CXCII. pp. 271—274, 1899. For other methods of dealing with problems involving many measurements see *Biometrika*, Vol. II. pp. 69—71, 77—78, 1902.

formation of tables, for the calculation of the means of classes, the standard deviation of classes, or the correlations between them, thus reducing the labour involved in the determination of intra-class or inter-class correlations to a small fraction of that hitherto required. While these short methods when properly applied give coefficients agreeing exactly with those deduced from tables, they cannot be recommended without two special warnings. (*a*) In the hands of inexperienced calculators abstract formulae are likely to result in blunders which would not occur, or at least could be easily detected, if the computations were made from tabled data. (*b*) The tables are, in some of the cases at least, desirable for other purposes, *e.g.*, the testing of linearity of regression.

The underlying principle is very simple. The familiar product moment formula for r_{xy} requires merely the first two moments of x and y around the means and the product moment $\Sigma(x'y')$. Of course, the rough moments may be taken around any origin whatever and referred to the mean by a proper formula. In many cases, there are material advantages* in taking 0 as the arbitrary origin. $\Sigma(x')$, $\Sigma(y')$ are then merely summations of the values of the individual variates; $\Sigma(x'^2)$, $\Sigma(y'^2)$ are summations of their squares; $\Sigma(x')/N$ is the mean, m and $\sigma_x = \sqrt{\Sigma(x'^2)N - m_x^2}$. Then

$$r_{xy} = \frac{\Sigma(x'y')/N - m_x m_y}{\sigma_x \sigma_y} \quad \text{.............................(i).}$$

Now the calculation of the rough moments and the product moments about 0 as origin for individual classes is a very easy task. Indeed, with the aid of a proper machine they may be rapidly obtained without seriating the measurements at all †. Since these rough moments for classes are taken about 0, they can be summed, weighted and summed, or subtracted in any way we please to obtain the rough moments for a correlation surface for the whole population. All of this is rapid machine work requiring but a moiety of time needed in the formation of tables of thousands of entries. Formula (i) is so well known that I shall do nothing further in the way of proof of the validity of the formulae suggested here than to show that the data from the first two moments and the numbers of individuals in the classes can be thrown into an equivalent form.

A word may be said concerning the illustrations given. All are, I believe, of intrinsic biological interest, but they have been chosen primarily for their convenient shortness, which facilitates not only the publication of the raw data, but the checking through of the arithmetical work by the reader. Thus they do not always represent the biologically best series which might have been selected.

* Harris, J. Arthur, "The Arithmetic of the Product Moment of Calculating the Coefficient of Correlation," *American Naturalist*, Vol. XLIV. pp. 693—699, 1910.

† After a little practice with the machine, one can sum the squares almost as readily as the first powers entered on the record sheets, providing the grades are not too high.

II. STATEMENT AND ILLUSTRATION OF PROBLEMS.

Problem I. To form Direct or Cross Intra-class or Inter-class Correlation or Contingency Tables when the Number of Possible Combinations is Large.

A method for the rapid formation of the complete direct intra-class table has been described and illustrated elsewhere *. The method of dealing with both direct and cross intra-class correlations by means of condensed tables has also been indicated†. Both direct and cross inter-class correlation tables (full or condensed) are formed by the same method, but more simply since the entries are $S(pq)$, not $S[n(n-1)]$, and the tables as first formed can be used without reduction. The method of dealing with fractional inter-class correlation tables will be illustrated as a check under Problem V.

Problem II. To Determine the Correlation between a Character and an Array of Associated Characters.

This serves chiefly as a simple introduction to the more complex cases to be described presently. One wishes often to correlate between a first measure and an array of associated measures‡. Tables may be difficult to form or cumbersome, while moments of the arrays of measurements around 0 as origin, $\Sigma(y')$ and $\Sigma(y'^2)$, may be wanted for other purposes, or, if not, are at least quickly calculated. Then, by the use of a machine for simultaneous multiplication and addition, we get for the rough moment coefficients

$$S(nx')/S(n) = \nu_x' = m_x, \quad S(nx'^2)/S(n) = \nu_x'' \quad \ldots\ldots\ldots\ldots\ldots(\text{ii}),$$

$$S(y')/S(n) = \nu_y' = m_y, \quad S\Sigma(y'^2)/S(n) = \nu_y'' \quad \ldots\ldots\ldots\ldots\ldots(\text{iii}),$$

and for the rough product moment coefficients

$$S\{x'[\Sigma(y')]\}/S(n) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(\text{iv}),$$

where n is the number of y measurements associated with any x, $N = S(n)$ is the number of measurements in all arrays or classes, Σ denotes a summation within the class, and S a summation of classes for the whole population. From this point the work is straightforward by (i), *i.e.*

$$r = \frac{S\{x'[\Sigma(y')]\}N - m_x m_y}{\sigma_x \sigma_y}.$$

The particular advantages of this method of calculating correlation are seen in cases where: (*a*) the class moments have already been calculated for other

* Harris, J. Arthur, "On the Formation of Correlation and Contingency Tables when the Number of Combinations is large," *American Naturalist*, Vol. XLV. pp. 566—571, 1911.

† Harris, J. Arthur, "The Formation of Condensed Correlation Tables when the Number of Combinations is large," *American Naturalist*, Vol. XLVI. pp. 477—486, 1912.

‡ A single parent may produce a score or a hundred offspring. Or, one may desire to know the relationship between the weight of the seed planted and the weight of the seeds produced, or between the number of pods per plant and the number of ovules and seeds produced per pod.

purposes, (*b*) it is desirable to calculate several constants from the same material with modified grouping or weighting of some of the classes, (*c*) it is desired to obtain the correlation without seriating the measures for the classes at all*.

Problem III. To Determine Direct and Cross Intra-class Coefficients of Correlation from the First Two Moments of the Individual Classes.

Let x be the measure of a first and y the measure of a second character of an individual of one of the several classes forming the population N. Then if correlations for the population be determined, r_{xy}, where x and y are measures upon the same individual, may be called an organic correlation, while $r_{x_1x_2}$, $r_{y_1y_2}$ may be designated as direct and $r_{x_1y_2}$ as a cross intra-class correlation, where the subscripts 1 and 2 indicate that the measures are taken on a "first" and a "second" individual of the class respectively.

A. *Direct Intra-class Relationships.*

Let n be the number of individual measurements, *e.g.* of siblings or homotypes, in one of m classes. Let Σ indicate a summation within any class and S a summation of classes. For an individual class the rough moment coefficients around 0 as origin are $\Sigma(x')/n$ and $\Sigma(x'^2)/n$, while the possible combinations of individuals are $n(n-1)$. Since the table is symmetrical $m_1 = m_2$, $\sigma_1 = \sigma_2$, and the mean product moment for any class is

$$\{[\Sigma(x')]^2 - \Sigma(x'^2)\}/n(n-1) \quad \text{..........................(v)},$$

or for the whole population of m classes

$$S\{[\Sigma(x')]^2 - \Sigma(x'^2)\}/S[n(n-1)] \quad \text{....................(vi)},$$

or when n is constant from class to class

$$S\{[\Sigma(x')]^2 - \Sigma(x'^2)\}/m[n(n-1)] \quad \text{....................(vii)}.$$

To complete the calculation of r only the first and second moments for the population are necessary. Two cases are possible: (*a*) the n's are identical for all classes, and (*b*) n differs from class to class. In case (*a*) the moment coefficients for the whole weighted population of $S[n(n-1)] = m[n(n-1)]$ individuals are unchanged by weighting, and are given by

$$S[\Sigma(x')]/N, \quad S[\Sigma(x'^2)]/N \quad \text{........................(viii)}.$$

For case (*b*) the arithmetical routine is a little longer, the classes being weighted differently, but always in a $(n-1)$-fold manner. Thus the moment coefficients are

$$S[(n-1)\Sigma(x')]/S[n(n-1)], \quad S[(n-1)\Sigma(x'^2)]/S[n(n-1)] \text{......(ix)}.$$

From this point the calculation of r is straightforward by (i). For clearness both case (*a*) and (*b*) will be illustrated.

* In this case the data for the means and standard deviations are obtained by summing first and second powers of the numbers from the record sheets without classification.

Illustration (III—Aa) 1. *Symmetry in the Ovary of Hibiscus.*

The morphologist interested in the degree of perfection of morphogenetic processes may desire to measure the degree of precision with which the different "*Anlagen*" which go to make up a composite organ, say for example a quinque-locular fruit, are differentiated off from the common primordium and carried through embryogenesis. One convenient measure is furnished by the intra-ovarial coefficient of correlation.

Take as an illustration a series of 1000 five-celled fruits of *Hibiscus Syriacus* from the Missouri Botanical Garden in the fall of 1905. Table I gives the symmetrical intra-ovarial correlation surface for number of ovules per locule.

Since all fruits have the same number of locules, the constants are not changed by weighting, hence, from *Biometrika* (Vol. VIII., pp. 61—62; M.B.G., 1905) we take $\bar{o}=6{\cdot}4648$, $\sigma_0={\cdot}892166$. From Table I,

$$S(o_1'o_2')/N = 841990/20000 = 42{\cdot}0995,$$

$$r=\{S(o_1'o_2')/N-\bar{o}^2\}/\sigma_0^2=\{42{\cdot}0995-(6{\cdot}4648)^2\}/({\cdot}892166)^2={\cdot}3843.$$

TABLE I. *Symmetry in the Ovary of Hibiscus.*

Ovules of Second Locule.

Ovules of First Locule.	2	3	4	5	6	7	8	Totals
2	—	—	1	2	6	2	1	12
3	—	—	—	1	7	4	8	20
4	1	—	12	28	132	56	23	252
5	2	1	28	220	899	255	55	1460
6	6	7	132	899	6384	1871	605	9904
7	2	4	56	255	1871	2014	1134	5336
8	1	8	23	55	605	1134	1190	3016
Totals	12	20	252	1460	9904	5336	3016	20000

But in *Biometrika* (*loc. cit.* Table VI) we have seriations of total ovules per fruit, $=\Sigma(o')$ for locules. From this and the distribution for individual locules we have for the product moments, by (vii),

$$S[\Sigma(o')]^2=1054938,\quad S(o'^2)=212948,\quad m[n(n-1)]=20000,$$

whence

$$r=\frac{\{S[\Sigma(o')]^2-S(o'^2)\}/[mn(n-1)]-\bar{o}^2}{\sigma_0^2}=\{42{\cdot}0995-(6{\cdot}4648)^2\}/({\cdot}892166)^2={\cdot}3843,$$

avoiding all the labour of forming a correlation surface of 20000 entries, but giving no information at all as to the nature of the regression curve.

Illustration (III—Aa) 2. *Intra-ovarial Correlation for Seed Production.*

Next consider a physiological problem from the same data. Table II gives the intra-ovarial (or inter-locular) correlation surface for seeds per locule.

$$\bar{s}=3{\cdot}9056,\quad \sigma_s=1{\cdot}756442,\quad S(s_1's_2')=339360,$$

whence $$r=\{S(s_1's_2')/20000-(3{\cdot}9056)^2\}/(1{\cdot}756442)^2={\cdot}5557.$$

TABLE II.

Intra-Ovarial Correlation for Seeds per Locule.

Seeds of Second Locule.

Seeds of First Locule.	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	Totals
0	70	123	126	54	31	12	4	—	—	420
1	123	334	414	311	150	72	47	8	1	1460
2	126	414	714	642	447	204	102	23	4	2676
3	54	311	642	1002	854	523	237	47	14	3684
4	31	150	447	854	1036	868	522	102	26	4036
5	12	72	204	523	868	1010	752	179	52	3672
6	4	47	102	237	522	752	916	252	84	2916
7	—	8	23	47	102	179	252	150	79	840
8	—	1	4	14	26	52	84	79	36	296
Totals	420	1460	2676	3684	4036	3672	2916	840	296	20000

But the same result can be obtained from the constants for seeds per locule and the seriation of total seeds per fruit (*Biometrika*, Vol. VIII., pp. 60—62) in precisely the same way as in the preceding illustration *without the excessive labour of tabling* 20000 *combinations.*

Suppose, now, we had been given the physical constants only—not the distributions—for seeds (or ovules) per locule and total seeds (or ovules) per fruit. The correlation is still easily deduced, for the sum of the squares of total seeds, $= S[\Sigma(s')]^2$ for locules, is given by substituting from *Biometrika*, Vol. VIII., Table XII,

$$S[\Sigma(s')]^2/N = \Sigma(s_t'^2)/N = \sigma_{s_t}^2 + \bar{s}_t^2 = (7{\cdot}0506)^2 + (5 \times 3{\cdot}9056)^2 = 431{\cdot}05374436,$$

where the subscript t indicates total seeds per fruit. Since $N = 1000$ for fruits,

$$S[\Sigma(s')]^2 = 431053{\cdot}74.$$

In the same manner we get for the individual locules

$$\Sigma(s'^2)/N = \sigma_s^2 + \bar{s}^2 = 18{\cdot}338652,$$

or since N for locules $= 5000$, $\Sigma(s'^2) = 91693{\cdot}26$. Hence, by (vii),

$$r = \{(431053{\cdot}74 - 91693{\cdot}26)/20000 - (3{\cdot}9056)^2\}/(1{\cdot}7564)^2$$

$$= 1{\cdot}71431264/3{\cdot}084941 = {\cdot}5557.$$

The extent of agreement with other methods depends merely upon the number of decimal places retained in the constants.

The importance of intra-class correlation coefficients in physiological work is emphasized by a comparison of this with the preceding, III (Aa) 1, result:

For ovules per locule, $r = {\cdot}3843 \pm {\cdot}0081$.
For seeds per locule, $r = {\cdot}5557 \pm {\cdot}0066$.
Excess for seeds, $= {\cdot}1714 \pm {\cdot}0104$.

The probable errors were calculated on the basis of the actual number of locules, $n = 5000$.

Apparently differences (a) in the vigour of the ovules, or (b) in the chances of fertilization, or (c) in the supply of seed forming substances, from fruit to fruit tend to render the locules more alike in the number of seeds they produce than they are initially in number of ovules*. The discussion of these points must, of course, be reserved until all the data can be set forth.

Illustration (III—Aa) 3. *Intra-ovarial Correlation for Seed Weight.*

Case (Aa) may be further illustrated by a series of seed weights to be discussed in a forthcoming paper on *Crinum.* Ten seeds from each of 200 fruits taken from a fine row of plants in the Missouri Botanical Garden in the spring of 1906 were weighed. From Illustration 2 it appeared that in *Hibiscus* the locules of the same fruit have a tendency to produce similar numbers of seeds. Here the problem is to determine to what extent the seeds from the same fruit are alike in weight. The seeds range from 1—18 half-gram units†. This would give a symmetrical intra-fruit correlation table of $18^2 = 324$ compartments and $200 \times 10 \times 9 = 18{,}000$ entries. The work is very rapidly carried out by my formulae. The total weight in units for each of the 200 fruits is obtained by simple addition. These, shown in Table III, are of course the class $\Sigma(x')$'s around

TABLE III.

Total weight of seeds (in half-gram units) in Individual Fruits of Crinum.

10	14	19	21	24	27	29	31	33	36	39	41	43	45	49	54	61	70	85	96
10	14	20	21	24	27	29	31	34	36	39	41	43	46	49	55	62	71	86	97
10	15	20	21	24	27	29	32	34	36	39	42	43	46	50	55	62	72	87	100
11	15	20	21	24	28	29	32	34	36	40	42	44	46	50	57	62	73	89	103
12	16	20	22	24	28	30	32	34	37	40	42	44	47	50	58	63	75	89	107
12	17	20	22	25	28	30	32	34	37	40	42	44	47	51	58	64	76	90	115
12	17	20	22	25	28	30	32	35	38	41	42	44	47	52	59	64	78	91	122
13	17	21	23	25	28	30	32	35	38	41	42	45	47	52	59	64	80	93	124
13	18	21	23	25	28	30	32	35	39	41	42	45	48	53	60	66	81	94	125
14	19	21	23	26	29	31	33	35	39	41	43	45	48	53	61	69	84	95	126

0 as origin. The 2000 seeds weighed are distributed as follows:

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
219	396	352	290	212	147	106	76	55	39	27	26	29	13	5	6	1	1

* This is also the conclusion reached for *Sanguinaria.* See Harris, *Biometrika*, Vol. VII. pp. 327—328, 1910.

† Class 1=0·0—0·5 gr., class 2=0·6—1·0 gr., etc. For details see final paper, *Annual Report Missouri Botanical Garden*, Vol. XXIII. 1912.

With Sheppard's correction this gives* in units of weighing

$$\Sigma(x') = 8778\dagger, \quad \Sigma(x'^2) = 56350,$$

$$\Sigma(x')/N = 4{\cdot}389, \quad \sigma^2 = 8{\cdot}828346.$$

From Table III $$S[\Sigma(x')]^2 = 510496,$$

and we have $$r = \frac{(510496 - 56350)/18000 - (4{\cdot}389)^2}{8{\cdot}828346} = {\cdot}6759,$$

agreeing exactly with the result from the conventional symmetrical table which will be published and discussed elsewhere.

Illustration (III—Ab) 4. Fraternal Correlation when the Number of Combinations is large.

As material for an illustration of the calculation of fraternal (or rather sororal) correlation when the numbers are large, Dr Lutz kindly gave me the measurements of wing lengths in units of $\frac{1}{33}$rd mm. for the females in fifteen families of the Pomace Fly, *Drosophila ampelophila*‡. Table IV shows the distributions of wing length, the class centres only being indicated. Using the method of a former note§, I formed the symmetrical intra-family correlation surface shown in Table V. From this I deduced for the weighted population

$$\bar{l} = 66{\cdot}200653, \quad \sigma_l = 7{\cdot}107332,$$

and for the first rough moments about 0 as origin of the individual arrays the numbers given in the last column of the table. Multiplying up by the grade of the array and summing simultaneously on the machine‖, I got

$$\Sigma(l_1 l_2) = 123391132,$$

whence $$r = [\Sigma(l_1' l_2')/N - \bar{l}^2]/\sigma_l^2 = {\cdot}15883.$$

But, entirely without the labour of forming a heavy table, I find, from the first and second class (family) moments of Table IV,

$$Sn(n-1) = 28148,$$

$$S[(n-1)\Sigma(l')] = 1863416,$$

$$S[(n-1)\Sigma(l'^2)] = 123559412,$$

$$S[\Sigma(l')]^2 = 126124937,$$

$$S\Sigma(l'^2) = 2733805.$$

* Of course, where all the classes comprise the same number of individuals the weighting is uniform and the weighted population constants the same as are those calculated from the unweighted distributions.

† Table III may be conveniently verified by the agreement of its summation with this number.

‡ See *Publications Carnegie Institution of Washington*, 143, pp. 38—40, for some details.

§ *American Naturalist*, Vol. XLV. pp. 566—571, 1911.

‖ *American Naturalist*, Vol. XLIV. pp. 693—699, 1910.

TABLE IV.

Wing Length in Families of Drosophila.

Family	*57*	*59*	*61*	*63*	*65*	*67*	*69*	*71*	*73*	Totals	$\Sigma(l')$	$\Sigma(l'^2)$
31	1	—	1	2	8	12	6	1	—	31	2053	136183
32	—	—	—	—	—	1	9	2	—	12	830	57420
37	1	1	—	3	5	15	9	4	4	42	2832	191426
39	—	1	2	5	10	12	12	7	1	50	3348	224634
57	—	—	—	4	6	18	9	—	—	37	2469	164877
58	—	—	—	1	16	16	11	4	—	48	3218	215928
523	—	—	1	3	7	15	19	8	1	54	3662	248654
575	—	—	3	5	4	3	3	—	—	18	1166	75658
576	—	—	6	12	16	7	8	5	—	54	3538	232270
582	—	—	2	10	12	21	6	2	—	53	3495	230749
587	—	—	—	—	—	2	12	4	—	18	1246	86274
588	—	—	—	4	17	20	12	—	—	53	3525	234613
1177	2	1	5	7	14	14	3	—	—	46	2974	192646
1193	—	—	2	12	18	16	5	—	—	53	3465	226749
1226	—	—	4	18	21	8	1	—	—	52	3348	215724
Totals	4	3	26	86	154	180	125	37	6	621	41169	2733805

TABLE V.

Sororal Correlation for Wing Length in the Vinegar Fly.

Wing of Second Sister.

Wing of First Sister.	*57*	*59*	*61*	*63*	*65*	*67*	*69*	*71*	*73*	Totals	$\Sigma(l_x')$
57	2	3	11	19	41	55	21	5	4	161	10605
59	3	—	7	15	29	41	24	11	5	135	8977
61	11	7	74	253	357	278	147	57	3	1187	77733
63	19	15	253	780	1234	1085	521	157	20	4084	268002
65	41	29	357	1234	2042	2128	1145	322	37	7335	483441
67	55	41	278	1085	2128	2558	1568	427	87	8227	545433
69	21	24	147	521	1145	1568	1212	440	67	5145	343643
71	5	11	57	157	322	427	440	158	31	1608	107682
73	4	5	3	20	37	87	67	31	12	266	17900
Totals	161	135	1187	4084	7335	8227	5145	1608	266	28148	1863416

Whence, writing the formula for r without reduction,

$$r_{l_1 l_2} = \frac{\dfrac{S[\Sigma(l')]^2 - S\Sigma(l'^2)}{S[n(n-1)]} - \left(\dfrac{S[(n-1)\Sigma(l')]}{S[n(n-1)]}\right)^2}{\dfrac{S[(n-1)\Sigma(l'^2)]}{S[n(n-1)]} - \left(\dfrac{S[(n-1)\Sigma(l')]}{S[n(n-1)]}\right)^2} = \cdot 15883,$$

agreeing exactly with the value deduced from the table.

While the case is given as an illustration of method only, chosen primarily because of the small number of families, it is of rather unusual biological interest. These families were from inbred stock (sister by brother) derived originally from a single pair, thus representing as nearly as possible among unisexual organisms a "pure line." Yet we find a sensible sororal correlation*. I hope Dr Lutz will find time to collect really adequate data on this interesting problem.

I illustrate on this data one of the most convenient features of the methods here described. If, for any reason, one wishes to compare the values of correlations determined with and without the inclusion of certain classes, the various determinations can be quickly made by the simple addition or subtraction of the summations of weighted products. For example, suppose one wishes to test the effect of omitting families 32, 575 and 587. $S[n(n-1)]$ will be reduced by

$$12 \times 11 + 2 \times 18 \times 17 = 744;$$

the first weighted population summation by

$$11 \times 830 + 17 \times 1166 + 17 \times 1246 = 50134;$$

the second by

$$11 \times 57420 + 17 \times 75658 + 17 \times 86274 = 3384464,$$

and the product summation by

$$(830^2 - 57420) + (1166^2 - 75658) + (1246^2 - 86274) = 3381620.$$

Deducting these from the comparable values given above for the population and calculating physical constants on the basis of the new weighted population N, we find $r = \cdot 1419$.

Illustration (*III—Ab*) 5. *Direct Homotypic Correlation.*

The records of the number of ovules per pod in twelve individual trees of black locust, *Robinia Pseud-acacia*†, may illustrate the calculation of homotypic correlations. The series is selected because of the wide variation in n. Naturally the smallness of the number of classes which renders the material convenient as an illustration of method, makes it of little value biologically. The seriations of numbers of ovules per pod and of numbers of seeds per pod appear in Tables VI and VII. The essentials for the calculation of r appear in Table VIII.

* This may not of course be due to heredity. The environment of the unlaid eggs or differences in the experimental conditions during the larval period may be its source. But Dr Lutz considers sensible environmental heterogeneity out of the question.

† The correlation between the number of ovules and number of seeds developing per pod has been discussed for the combined pods of this series in *Biometrika*, Vol. VI. pp. 441—442, 1909.

TABLE VI.

Ovules per Pod in Robinia.

Tree	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	Totals
1	1	3	12	24	21	28	20	6	7	—	—	—	—	—	122
2	—	—	3	4	6	10	11	11	6	6	4	1	2	—	64
3	1	2	3	13	35	25	20	11	1	—	—	—	—	—	111
4	—	3	5	16	31	25	14	6	2	—	—	—	—	—	102
5	—	—	3	13	25	26	31	15	5	1	2	1	—	—	122
6	—	2	3	14	32	31	25	10	1	2	—	—	—	—	120
7	—	—	—	6	10	15	35	25	16	9	1	3	—	—	120
8	—	—	—	—	2	5	22	28	35	31	25	6	4	1	159
9	—	—	—	2	11	29	28	32	14	9	2	—	1	—	128
10	—	—	1	3	5	22	15	20	5	4	2	1	—	—	78
11	—	1	—	8	19	17	20	26	7	3	2	—	2	—	105
12	—	—	—	—	2	12	17	25	36	34	33	20	11	6	196
Totals	2	11	30	103	199	245	258	215	135	99	71	32	20	7	1427

TABLE VII.

Seeds per Pod in Robinia.

Tree	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	Totals
1	—	—	5	16	23	27	17	16	8	3	3	—	4	—	—	—	—	—	122
2	—	1	4	4	7	6	8	6	5	6	3	2	4	4	2	1	1	—	64
3	2	13	28	23	11	20	9	3	2	—	—	—	—	—	—	—	—	—	111
4	4	13	19	20	11	11	12	3	5	2	1	1	—	—	—	—	—	—	102
5	—	5	18	10	20	21	21	7	7	4	6	2	—	—	—	1	—	—	122
6	13	33	19	21	15	6	8	1	1	2	1	—	—	—	—	—	—	—	120
7	3	11	20	24	21	11	14	6	7	2	—	1	—	—	—	—	—	—	120
8	—	—	—	6	23	14	21	14	25	20	18	9	5	2	2	—	—	—	159
9	3	15	17	15	13	22	13	9	2	10	3	2	3	1	—	—	—	—	128
10	3	9	15	14	11	8	9	3	3	—	—	1	1	—	1	—	—	—	78
11	—	—	1	3	6	9	12	22	20	17	11	3	1	—	—	—	—	—	105
12	1	—	6	5	7	7	11	8	12	22	23	23	20	16	17	10	6	2	196
Totals	29	100	152	161	168	162	155	98	97	88	69	44	38	23	22	12	7	2	1427

Feeding the entries in Table VIII into the calculating machine and copying off the results I find, for population constants weighted in $(n-1)$-fold manner:

$$S[n(n-1)] = 181152,$$

$$S[(n-1)\,\Sigma(o')] = 2259669, \qquad \bar{o} = 12{\cdot}473884,$$

$$S[(n-1)\,\Sigma(o'^2)] = 29203787, \qquad \sigma_0^2 = 5{\cdot}613725,$$

$$S[\Sigma(s')]^2 = 13114598, \qquad S[\Sigma(s'^2)] = 101750,$$

$$S[\Sigma(o')]^2 = 28866082, \qquad S[\Sigma(o'^2)] = 219072,$$

$$S[(n-1)\,\Sigma(s')] = 1468970, \qquad s = 8{\cdot}109047,$$

$$S[(n-1)\,\Sigma(s'^2)] = 14364330, \qquad \sigma_s^2 = 13{\cdot}537706,$$

where o and s indicate ovules per pod and seeds per pod.

These give:

$$r_{oo_2} = \frac{\{S[\Sigma(o')]^2 - S[\Sigma(o'^2)]\}/S[n(n-1)] - \bar{o}^2}{\sigma_o^2} = \cdot 452,$$

$$r_{s_1 s_2} = \frac{\{S[\Sigma(s')]^2 - S[\Sigma(s'^2)]\}/S[n(n-1)] - \bar{s}^2}{\sigma_s^2} = \cdot 449.$$

The number of individuals is too small to attach much importance to the result but the agreement with other series of Leguminosae is good*, although the correlation for seeds maturing is frequently considerably smaller than that for ovules formed.

TABLE VIII.

Summations for Individuals of Robinia.

Number of Tree	n	$n(n-1)$	Sum of Ovules, $\Sigma(o')$	Sum of Ovules Squared, $\Sigma(o'^2)$	Sum of Seed, $\Sigma(s')$	Sum of Seeds Squared, $\Sigma(s'^2)$
1	122	14762	1273	13649	912	7394
2	64	4032	796	10256	602	6522
3	111	12210	1183	12839	595	3543
4	102	10302	1064	11310	586	3938
5	122	14762	1378	15896	864	6918
6	120	14280	1299	14311	555	3091
7	120	14280	1490	18850	713	4819
8	159	25122	2251	32331	1495	15117
9	128	16256	1580	19816	860	6894
10	78	6006	946	11690	457	3211
11	105	10920	1245	15135	984	9642
12	196	38220	2875	42989	2347	30661
Totals	1427	181152	17380	219072	10970	101750

Illustration (III—Ab) 6. Tests of the Existence of Sub-races.

O. H. Latter in testing for differentiation in Cuckoo's eggs† according to their foster parents, determined the probable errors of the difference between the means of the several "gentes" and the means of *Cuculus Canorus* as a species. But one may also test for the differentiation of the foster parent "gentes" from the species as a whole by determining the intra-class correlation for egg dimensions. I am especially willing to do this since Latter tested the significance of his divergences by the formula $\cdot 67449\sqrt{\frac{\sigma^2}{n}+\frac{\Sigma^2}{N}}$, whereas Pearson's formula

$$\cdot 67449\sqrt{\frac{\Sigma^2}{N}+\frac{\sigma^2}{n}\left(1-\frac{2n}{N}\right)-\frac{n(M-m)^2}{N(N-n)}}$$

would have been more suitable but had not then been published‡.

* Pearson, *Phil. Trans.* A, Vol. CXCVII. pp. 364—379; Harris, *American Naturalist*, Vol. XLV. p. 570.

† *Biometrika*, Vol. I. pp. 164—176, 1901; Vol. IV. pp. 363—373, 1905.

‡ *Ibid.*, Vol. V. pp. 181—183, 1906.

Table III of Latter's second memoir furnishes the data. Grouping in units of ·5 mm. we have the results given in Tables IX and X*.

TABLE IX.

Length of Cuckoo's Eggs.

Foster Parent	*38*	*39*	*40*	*41*	*42*	*43*	*44*	*45*	*46*	*47*	*48*	*49*	*50*	*51*	*53*	Totals	Σ (l′)	Σ (l′²)
Robin	—	—	1	1	8	3	9	13	20	6	11	2	2	1	1	78	3559	162819
Garden-Warbler...	—	—	2	4	15	15	19	25	10	—	1	—	—	—	—	91	3988	174992
Meadow-Pipit ...	—	1	1	2	4	10	27	19	8	7	5	2	—	—	—	86	3837	171489
Wren	—	—	7	5	14	8	9	6	3	2	—	—	—	—	—	54	2315	99433
Redstart	—	—	1	—	5	6	21	8	11	6	5	1	—	—	—	64	2869	128821
Sedge-Warbler ...	—	—	2	—	4	11	15	5	6	3	1	1	—	—	—	48	2120	93788
Barred-Warbler ...	—	—	1	4	4	4	11	7	3	—	1	—	—	—	—	35	1529	66899
Whitethroat ...	1	—	—	3	7	6	8	8	2	—	2	—	—	—	—	37	1613	70457
Marsh-Warbler ...	—	—	1	—	3	4	14	6	7	1	1	—	—	—	—	37	1641	72867
Red-backed Shrike	—	—	2	9	20	24	90	67	60	23	9	3	—	—	—	307	13716	613606
White-Wagtail ...	—	—	—	2	10	7	29	33	23	6	4	1	—	—	—	115	5145	230453
Pied-Wagtail ...	—	—	2	1	7	9	20	13	22	5	5	—	1	—	—	85	3804	170532
Reed-Warbler ...	—	—	2	—	7	9	13	7	13	9	4	2	—	—	—	66	2959	132939
Tree-Pipit ...	—	—	—	—	6	2	11	7	6	7	9	—	—	—	—	48	2174	98648
Hedge-Sparrow ...	—	—	—	—	2	5	14	13	13	3	5	—	3	—	—	58	2629	119357
Yellow-Ammer ...	—	—	1	—	6	3	5	7	5	3	—	—	1	—	—	31	1377	61293
Totals ...	1	1	23	31	122	126	315	244	212	81	63	12	7	1	1	1240	55275	2468393

TABLE X.

Breadth of Cuckoo's Eggs.

Foster Parent	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	*38*	Totals	Σ (b′)	Σ (b′²)
Robin	—	—	1	2	17	17	20	14	6	1	78	2621	88233
Garden-Warbler ...	—	—	—	8	39	20	21	3	—	—	91	2975	97355
Meadow-Pipit ...	—	—	1	4	18	19	36	6	2	—	86	2863	95425
Wren	—	—	3	11	19	7	10	4	—	—	54	1750	56810
Redstart	—	—	2	2	17	19	22	2	—	—	64	2111	69703
Sedge-Warbler ...	—	—	—	8	11	13	12	4	—	—	48	1577	51881
Barred-Warbler ...	—	—	2	5	10	9	8	1	—	—	35	1139	37119
Whitethroat ...	1	—	—	5	13	11	6	1	—	—	37	1201	39041
Marsh-Warbler ...	—	—	—	3	14	11	9	—	—	—	37	1210	39602
Red-backed Shrike	—	1	2	10	80	104	94	14	2	—	307	10149	335833
White-Wagtail ...	—	—	3	5	20	27	32	21	7	—	115	3851	129177
Pied-Wagtail ...	—	—	—	5	27	21	28	4	—	—	85	2804	92590
Reed-Warbler ...	—	—	—	4	19	13	20	9	1	—	66	2192	72898
Tree-Pipit ...	—	—	—	2	9	17	14	3	3	—	48	1600	53398
Hedge-Sparrow ...	—	—	—	1	9	18	20	9	1	—	58	1944	65220
Yellow-Ammer ...	—	—	—	1	12	7	10	1	—	—	31	1021	33657
Totals ...	1	1	14	76	334	333	362	96	22	1	1240	41008	1357942

* Since in the method of computation adopted we deal with the summations of first and second powers of actual measurements, not with their deviations from arbitrary origins chosen to lighten computation, we can save materially in the size of the summed products by reducing to half mm. units instead of mm. units, i.e. 15·0 = 30, 15·5 = 31, etc.

Summarizing the results from these tables as under illustration (III—Ab) 5, we have:

$$S[n(n-1)] = 159484,$$

$$\begin{array}{llll} S[(n-1)\Sigma(l')] &= 7120158, & \bar{l} &= 44\cdot644967, \\ S[(n-1)\Sigma(l'^2)] &= 318370242, & \sigma_l^2 &= 3\cdot078842, \\ S[\Sigma(l')]^2 &= 320370295, & S[\Sigma(l'^2)] &= 2468393, \\ S[(n-1)\Sigma(b')] &= 5277568, & \bar{b} &= 33\cdot091520, \\ S[(n-1)\Sigma(b'^2)] &= 174844798, & \sigma_b^2 &= 1\cdot266907, \\ S[\Sigma(b')]^2 &= 176007866, & S[\Sigma(b'^2)] &= 1357942. \end{array}$$

Whence by (vi, ix) as illustrated in greater detail under (III—Ab) 4—5 above I get:

$$r_{l_1l_2} = \cdot046 \pm \cdot019, \qquad r/E_r = 2\cdot41,$$

$$r_{b_1b_2} = \cdot036 \pm \cdot019, \qquad r/E_r = 1\cdot85,$$

where the probable errors are calculated on the basis of the actual number of eggs, or $N = 1240$.

Such a test as the above can never replace the *R.L.D.* (relative local difference) or the *R.R.D.* (relative racial difference) as the case may be, of Tocher and Pearson. But there are cases, as for instance the present and several unpublished series, where, now that a (relatively) rapid method of computation is available, it may be very useful.

Illustration (III—Ab) 7. Inheritance of Milk Yield in the Male Line.

The test of the inheritance of a character like "milk yield" through the male must be made either by correlating between the milk yield of a cow and her parental granddam, or by determining the correlation between the daughters of individual bulls.

The methods here described are particularly useful for cases of this kind, for while the number of daughters in a family may not be large, the range of variation in yield is apt to be very great thus rendering the tables cumbersome in size, whether or not heavy in number of entries.

As an illustration I take a series of data from a paper by James Wilson, M.A., B.Sc., "On inheritance of Milk Yield in Cattle*." One of his tables (pp. 103—105), gives the milking record of the daughters of eleven Red Danish bulls. Our problem is to determine the correlation in milk production between the daughters of the same sire, in order to ascertain whether there be an inheritance of milking quality in the male line.

The data are of course very rough since they are apparently for the most part recorded to the nearest fifty gallon point. The yields range from 650 to 1400 gallons. Omitting all records indicated as questionable, I sum the yields and

* Wilson, J., *Scientific Proc. Roy. Dublin Society*, n. s. Vol. XIII. pp. 89—113, 1911.

the squares of yields directly from Professor Wilson's table and enter the totals for each bull in Table XI. This gives

$$r_{y_1y_2} = \cdot 126 \pm \cdot 076.$$

The probable error is based on the actual number of daughters, $n = 76$, not the weighted number,

$$S[n(n-1)] = 538.$$

Whether or not this slight correlation in yield (y) is statistically significant, the case illustrates a very convenient method of dealing with such problems. I hope to discuss the biological sides of the question in more detail elsewhere.

TABLE XI.

Milk Yield in Daughters of Various Bulls.

Sire	Number of Daughters	Total Yield $\Sigma(y')$	$(n-1)\,\Sigma(y')$	$\Sigma(y^2)$	$(n-1)\,\Sigma(y^2)$
Maxi	11	8900	89000	7310000	73100000
Stamfadern ...	9	8000	64000	7195000	57560000
Braendekilde Max	4	3800	11400	3625000	10875000
Vigfus	6	5900	29500	6090000	30450000
Osvald Ejersminde	6	5500	27500	5120000	25600000
Gunnar	3	2650	5300	2362500	4725000
Mazeppa II ...	6	6400	32000	6875000	34375000
Tordensjold ...	5	4650	18600	4372500	17490000
Trym I	7	6500	39000	6165000	36990000
Taurus IV ...	13	11350	136200	10022500	120270000
Ambrosius III ...	6	5300	26500	4765000	23825000
	76	68950	479000	63902500	435260000

$S[n(n-1)] = 538$, $\quad S[\text{Total yield}]^2 = 491477500$.

B. *Cross Intra-Class Relationships.*

If x and y be two measures on the same individual, and if correlation be determined between the individuals of the m sub-groups or classes constituting the general population, then as pointed out above, $r_{x_1x_2}$, $r_{y_1y_2}$ are direct intra-class correlations while $r_{x_1y_2}$ is a cross intra-class relationship.

The number of combinations will in both direct and cross intra-class relationships be $S[n(n-1)]$, since in the direct correlations one cannot correlate between a measure and itself and in the cross correlations one must not include the relationship between the x and y of the same individual—the organic correlation, r_{xy}.

The value of this direct organic correlation is given by the moment coefficients

$$S[\Sigma(x')]/S(n), \quad S[\Sigma(x'^2)]/S(n), \quad S[\Sigma(y')]/S(n), \quad S[\Sigma(y'^2)]/S(n), \ldots \text{(x)}$$

and by the mean product moment

$$S[\Sigma(x'y')]/S(n), \qquad \text{(xi)}$$

taken about 0 as origin. Most simply they can be determined by the formation of an ordinary organic correlation table for the whole population which will then be described by N, $S(x')$, $S(x'^2)$, $S(y')$, $S(y'^2)$ and $S(x'y')$, where S denotes a summation for the individuals of the population. This correlation will generally be wanted for its own sake.

Again two cases are possible (α) all the m classes are equally large, (β) n is variable. The first two moments are given by the same formulae as in the direct intra-class correlation above, (viii)—(ix). The mean product moment around 0 is

$$\{S[\Sigma(x')\Sigma(y')] - S[\Sigma(x'y')]\}/S[n(n-1)], \qquad \text{(xii)}$$

or with a denominator of $m[n(n-1)]$ where the m classes are equally large.

Illustration (*III—Ba*) 8. *Cross inter-locular* (*intra-ovarial*) *Correlation in Hibiscus.*

For a purpose quite foreign to my present one, the correlation between total ovules per fruit ($=\Sigma(o')$ for locules) and total seeds per fruit ($=\Sigma(s')$ for locules) has been tabled for the 1905 series of *H. Syriacus*. The Table has 17×36 entries, hence we reduce to the condensed form in Table XII. The organic correlation between the ovules and seeds of the same locule is shown in Table XIII. Taking the product moments of these two tables about 0, and subtracting that for locules from that for fruits to remove the products of the ovules and seeds of the same locule, and to reduce N from $S(n^2) = 25000$ to $S[n(n-1)] = 20000$, I find

$$S[\Sigma(o')\Sigma(s')] = 637491, \quad S(o's') = 128377,$$

and, substituting physical constants from illustrations 1 and 2, I find

$$r_{o_1 s_2} = \frac{S\{[\Sigma(o')\Sigma(s')] - S(o's')\}/20000 - \bar{o}\bar{s}}{\sigma_o \sigma_s} = \frac{25\cdot4557 - 6\cdot4648 \times 3\cdot9056}{\cdot892166 \times 1\cdot756442} = \cdot1320\,.$$

TABLE XII.

Correlation for Total Ovules and Total Seeds per Fruit in Hibiscus.

Ovules	f	Total Seeds	Ovules	f	Total Seeds
23	—	—	32	81	1512
24	3	41	33	80	1575
25	2	26	34	80	1436
26	4	51	35	85	1746
27	18	285	36	54	1256
28	45	708	37	54	1129
29	107	1847	38	34	807
30	188	3727	39	27	691
31	125	2322	40	13	369

TABLE XIII.

Correlation for Ovules and Seeds per Locule in Hibiscus.

Seeds per Locule.

Ovules per Locule.	0	1	2	3	4	5	6	7	8	Totals
2	2	1	—	—	—	—	—	—	—	3
3	—	3	2	—	—	—	—	—	—	5
4	4	8	28	17	6	—	—	—	—	63
5	6	47	76	109	86	41	—	—	—	365
6	49	169	315	490	573	525	355	—	—	2476
7	30	97	167	221	237	244	235	103	—	1334
8	14	40	81	84	107	108	139	107	74	754
Totals	105	365	669	921	1009	918	729	210	74	5000

Comparing other constants for Hibiscus:—

Direct Intra-class for Ovules, $r = \cdot3843 \pm \cdot0081$,

Direct Intra-class for Seeds, $r = \cdot5557 \pm \cdot0066$,

Cross Intra-class for Ovules and Seeds, $r = \cdot1320 \pm \cdot0093$,

Organic for Ovules and Seeds, $r = \cdot2722 \pm \cdot0088$.

In all cases the probable error is calculated on the basis of the actual, not the weighted number of locules.

Illustration (*III—Bb*) 9. *Cross Homotypic Correlation for Number of Ovules and Number of Seeds per Pod in* Robinia.

From Tables VI, VII, and VIII *supra*, we deduce $S[\Sigma(o')\,\Sigma(s')] = 19070536$ which taken in conjunction with $S(o's') = S[\Sigma(o's')] = 141476$ taken from the organic correlation table for ovules and seeds in *Biometrika* (Vol. VI. pp. 441—442) and the physical constants $\bar{o}$, $\bar{s}$, σ_o, σ_s, given under illustration (III—Ab) 5, gives

$$r_{o_1s_2} = \{\{S[\Sigma(o')\,\Sigma(s')] - S[\Sigma(o's')]\}/S[n(n-1)] - \bar{o}\bar{s}\}/\sigma_0\sigma_s = \cdot383.$$

This may be compared with $r_{o_1o_2} = \cdot452$ and $r_{s_1s_2} = \cdot449$. The discussion of the biological significance of the results from this and the preceding illustration must be reserved.

Problem IV. To Determine Direct or Cross Inter-Class Correlations from the First Two Moments of the Individual Classes.

Let x be the measure of a character in an individual of one class of p members and y be the measure of the same or a different character in an individual of another class of q members associated with the first for some logical (*e.g.*, biological, or sociological) reason. The coefficient calculated from a table in which every x of a first class is compared with every y in the second (associated) class may be designated as an inter-class coefficient of correlation.

The number of entries in an inter-class correlation surface will be $S(pq)$, where S indicates a summation of classes. $\Sigma(x')/p$, $\Sigma(x'^2)/p$, $\Sigma(y')/q$, $\Sigma(y'^2)/q$ are the

first two moment coefficients of individual classes, $[\Sigma(x')\Sigma(y')]/pq$ the product moment coefficient for an associated pair of classes and

$$S[\Sigma(x')\Sigma(y')]/S(pq), \quad \text{.........................(xiii)}$$

the rough product moment coefficient for the whole population.

For every first class of p individuals there is an associated second class of q individuals. Hence, the individuals of the first class are always weighted* in a q-fold manner and those of the second class in a p-fold manner. The population constants weighted for the calculation of r are given by the rough moment coefficients about 0 as origin,

$$S[q\Sigma(x')]/S(pq), \quad S[q\Sigma(x'^2)]/S(pq), \quad \text{..................(xiv)}$$

$$S[p\Sigma(y')]/S(pq), \quad S[p\Sigma(y'^2)]/S(pq). \quad \text{..................(xv)}$$

Since x and y may be measures of the same or different characters it is clear that the formula is the same for direct or cross inter-class correlation.

Illustration (*IV*) 10. *Racial Heterogeneity and Homogamy in Man.*

It has sometimes been suggested that homogamy for physical and mental characters in man is due to (*a*) the existence in the population of numerous local races, (*b*) the excessive frequency of intra-racial marriages, attributable to propinquity merely, and (*c*) the pooling of records from many local races in the same correlation surface.

A study of the literature of assortative mating convinces one that so simple an explanation is inadequate. Nevertheless, it is very important to give full weight to any factors tending to bring about a spurious correlation for assortative mating.

One way of testing the question is to make random matings within the local races, pool the results for a series of districts, and determine the correlation†. The influence of the local differentiation can be read off in the inter-sexual correlation coefficient thus secured‡.

A convenient series of materials for stature is that gathered by Tocher§ for the counties of Scotland. To be sure, his measurements are on individuals who have become insane, but data for the normal population are lacking.

We may work from the table of constants (*Biometrika*, Vol. v. Supplement, Table XIX), taking populations of arbitrary size, say 100 males, m, and 100

* When it is possible to collect data so that either p or q, or both, are constant throughout the m pairs or classes the work is obviously considerably simplified. Unfortunately such cases are rare in practice.

† This was actually done for parish churchyards in the cooperative study of assortative mating for duration of life. Cf. *Biometrika*, Vol. II. p. 485, 1903.

‡ Of course our problem has some similarity to but differs fundamentally from that considered by Tschepourkowsky in his study of the resemblance between man and woman. Cf. *Biometrika*, Vol. IV. pp. 161—168, 1905.

§ Tocher, J. F., "The Anthropometric Characteristics of the Inmates of Asylums in Scotland," *Biometrika*, Vol. v. pp. 298—350, 1907.

females, f, from each district. Then for a district, $\Sigma(s') = 100\,\bar{s}$, the population means are given by $S[\Sigma(s')]/N$, the population standard deviations* by

$$\text{S.D.}^2 = S[100(\bar{s}^2 + \sigma_s^2)]/N - [S(100\,\bar{s})/N]^2,$$

where s represents stature in males or females as the case may be, the bars and σ's indicate county means and standard deviations and $N = S(n) = mn = 22 \times 100$.

The population constants are:

For Males:

$$S(100\,\bar{s}) = 145080, \qquad N = 2200,$$
$$S(100\,\bar{s}^2) = 9567888, \qquad \bar{s} = 65{\cdot}945,$$
$$S(100\,\sigma_s^2) = 17335{\cdot}08, \qquad \text{S.D.}^2 = 8{\cdot}116546,$$

For Females:

$$S(100\,\bar{s}) = 134760, \qquad N = 2200,$$
$$S(100\,\bar{s}^2) = 8255232, \qquad \bar{s} = 61{\cdot}255,$$
$$S(100\,\sigma_s^2) = 13992{\cdot}05, \qquad \text{S.D.}^2 = 6{\cdot}618921.$$

These will suffice for getting the correlations, since all classes comprise the same number of individuals and no change in the $\bar{s}$ or σ_s is introduced by weighting.

The totals of the correlation surface are $22 \times 100 \times 99 = 217800$ for the intra-patrial correlation for stature of random pairs of males and random pairs of females, and the rough product moment coefficients are given by

$$\{100\,S(100\,\bar{s}^2) - [S(100\bar{s}^2) + S(100\,\sigma_s^2)]\}/S[n(n-1)],$$

whence, for males, $r = {\cdot}019$,

for females, $r = {\cdot}029$.

But for the inter-sexual correlation between random pairs of men and women for the several counties, pooled together, there are mn^2, not $m[n(n-1)]$ combinations, since each male may be compared with all the females of the district, while in intra-sexual correlations the individual's own measure must be deducted. For random pairs of men and women we have, therefore, for the unadjusted product moment

$$S[\Sigma(s'_m)\,\Sigma(s'_f)] = 888716500, \qquad N = 220000,$$

whence

$$r = (4039{\cdot}620454 - 65{\cdot}945455 \times 61{\cdot}254545)/2{\cdot}848955 \times 2{\cdot}572726 = {\cdot}022.$$

The result just deduced taken in comparison with that secured in a similar manner for the several parishes of the Wensleydale district†, seems to indicate that the geographical differentiation in the population plays a very small part in the production of the correlation between husband and wife.

* In getting $S[100(\bar{s}^2 + \sigma_s^2)]$ for the standard deviations we can save the labour of mental additions of large squares by taking separately $S(100\bar{s}^2)$ and $S(100\sigma_s^2)$. We have then only to multiply $S(100s^2)$ by 100 to get $S[100s]^2 = S[\Sigma(s')]^2$. $S\Sigma(s'^2)$ is of course given by the second moment calculated for the population S.D. Thus the squaring and summation can be quickly done simultaneously on any of the usual types of calculators.

† *Biometrika*, Vol. II. p. 485.

Problem V. To Determine Fractional Direct or Cross Inter-class Correlations from the First Two Moments of the Individual Classes.

Under (IV) every x of the first class is compared with every y of the second class. In actual practice, however, cases are frequent where all the possible comparisons are not made. Such relationships may be designated as fractional inter-class correlations. They are more difficult to determine than the simple inter-class constants but they are of very great importance in many cases. Of course the method is quite general, but for clearness I take the determination of the avuncular correlations as an illustration.

If x be the measure of a character in a parental array, family or sibship of p individuals and y the measure of the same or a different character in the offspring array (or family or families) of q individuals, then the inter-class correlation involving the comparison of every x with every y in the related classes of the population of m classes may be designated as an ascendant (or, if one chooses, descendant) correlation of the first order*. From such a table, actually formed by the method described under (I) or only expressed algebraically as under (IV), one may easily derive the avuncular correlation—*i.e.*, that between "uncles (aunts)" and "nephews (nieces)"—as follows:

In such an ascendant correlation, the offspring individuals are associated with all the members of the parental array, both their own parents and their parent's siblings, or there are $S(pq)$ combinations.

Hence it is given by formulae (xiii—xv) under (IV). To obtain the avuncular correlation we have only to deduct from the five fundamental summations of the ascendant correlation those for the relationships between (weighted) individual parents and their $S(q)$ individual offspring, and proceed to the correlation of r on the basis of $S(pq-q)$ instead of $S(pq)$ individuals.

But the direct parental correlation is practically always wanted on its own account. It is most easily obtained from an ordinary or condensed correlation table, or by the method described under (II). Designating by the subscript i the measures entering into the correlations for individual parents (weighted with their offspring) and their offspring, we have for the avuncular relationship (the fractional inter-class correlation) by simple subtraction,

$$\frac{S[\Sigma(x')\,\Sigma(y')] - S(x_i'y_i')}{S(pq) - S(q)}, \quad \text{.........................(xvi)}$$

for the product moment coefficient, and

$$\frac{S[q\Sigma(x')] - S(x_i')}{S(pq) - S(q)}, \qquad \frac{S[q\Sigma(x'^2)] - S(x_i'^2)}{S(pq) - S(q)}, \quad \text{............(xvii)}$$

$$\frac{S[p\Sigma(y')] - S(y_i')}{S(pq) - S(q)}, \qquad \frac{S[p\Sigma(y'^2)] - S(y_i'^2)}{S(pq) - S(q)}, \quad \text{............(xviii)}$$

for the moment coefficients.

* If correlation between parental and offspring arrays be designated as of the first order, that between grandparental and grandoffspring arrays is of the second order, and so on.

TABLE XIV. *Weight of Seeds produced by Parental and Offspring Plants in Garden Beans.*

Weight for Parental Plants

Plant	3	4	5	6	7	8	9	10	11	12	13	14	p	$\Sigma(x')$	$\Sigma(x'^2)$
1	1	–	–	–	–	5	10	6	1	–	–	–	23	204	1860
2	–	–	2	–	4	12	15	3	–	–	–	–	36	299	2529
3	–	–	–	–	1	3	10	9	2	–	–	–	25	233	2193
4	–	–	–	1	3	6	8	1	–	–	–	–	19	157	1315
5	–	1	–	2	–	2	6	21	10	–	–	–	42	406	4012
6	–	–	–	–	2	13	26	7	–	–	–	–	48	422	3736
7	–	–	–	–	–	2	8	15	15	–	–	–	40	403	4091
8	–	–	–	–	7	18	15	2	–	–	–	–	42	348	2910
9	–	–	–	1	–	5	9	13	12	5	–	–	45	449	4557
10	–	–	–	1	–	11	17	11	–	–	–	–	40	357	3217
11	–	–	–	–	–	3	7	7	13	2	–	–	32	324	3320
12	–	–	–	–	2	9	20	1	–	–	–	–	32	276	2394
13	–	–	1	10	18	8	1	–	–	–	–	–	38	264	1860
14	–	–	–	–	–	6	8	22	4	–	–	–	40	384	3716
15	–	–	–	1	2	6	12	17	2	–	–	–	40	368	3432
16	–	–	–	–	–	2	–	1	12	21	6	1	43	502	5914
17	–	–	1	1	5	11	2	3	–	–	–	–	23	182	1472
18	–	–	–	–	–	–	1	10	3	–	–	–	14	142	1444
19	–	–	–	2	2	14	5	–	–	–	–	–	23	183	1471
20	–	–	–	–	–	1	4	4	9	4	–	–	22	231	2453
21	–	–	–	3	5	9	6	3	2	–	–	–	28	231	1957
22	–	–	–	–	2	12	16	5	4	–	–	–	39	348	3146
23	–	–	–	–	3	11	16	1	–	–	–	–	31	263	2247
24	–	–	–	–	1	6	17	12	2	–	–	–	38	350	3252
25	–	–	–	–	–	1	19	20	1	–	–	–	41	390	3724
26	–	–	–	1	–	3	8	15	8	3	1	–	39	389	3945
27	–	–	–	–	4	14	19	4	–	–	–	–	41	351	3031
28	–	–	–	3	9	10	14	2	3	–	–	–	41	340	2886
29	–	–	2	1	1	7	14	13	3	–	–	–	41	368	3380
30	–	–	–	–	2	27	11	1	–	–	–	–	41	339	2817
31	–	–	–	–	–	1	3	19	13	4	–	–	40	416	4356
32	–	–	–	1	–	7	8	14	8	2	–	–	40	386	3788
33	–	–	–	–	1	3	15	23	2	–	–	–	44	418	3998
34	–	–	–	–	–	1	7	25	6	2	–	–	41	411	4145
35	–	–	–	–	–	1	–	13	24	2	1	–	41	439	4725
36	–	–	–	–	1	13	22	5	–	–	–	–	41	359	3163
37	–	–	–	–	1	–	6	4	13	7	–	–	31	328	3516
38	–	–	–	–	3	4	8	6	4	1	–	–	26	241	2279
39	–	–	1	–	–	4	19	12	6	–	–	–	42	394	3746
40	–	–	–	–	–	–	4	16	14	3	–	–	37	386	4050

Weight for Offspring Plants

Line	3	4	5	6	7	8	9	10	11	12	13	14	15	q	$\Sigma(y')$	$\Sigma(y'^2)$
1	–	–	–	2	2	9	15	5	1	1	–	–	–	35	306	2726
2	–	–	–	2	3	11	13	12	2	–	–	–	–	43	380	3418
3	–	1	2	8	7	11	5	2	2	–	–	–	–	38	286	2248
4	–	–	2	5	10	17	10	6	–	–	–	–	–	50	396	3218
5	–	–	2	4	4	8	12	5	4	1	–	–	–	40	340	3002
6	–	–	6	10	13	18	10	3	–	–	–	–	–	60	445	3409
7	–	–	1	2	5	12	18	5	–	1	–	–	–	44	372	3212
8	1	–	1	1	7	10	9	9	2	3	–	–	–	43	372	3356
9	1	3	–	1	2	5	11	9	2	–	–	–	–	34	286	2544
10	–	–	1	1	2	5	10	5	5	–	–	–	–	29	260	2394
11	–	–	–	6	12	11	9	5	1	1	–	–	–	45	362	3002
12	–	–	3	6	7	15	8	4	1	–	–	–	–	44	343	2763
13	–	–	2	8	14	8	12	1	1	–	–	–	–	46	349	2729
14	–	1	3	4	13	10	4	5	–	–	–	–	–	40	300	2336
15	–	1	1	5	5	10	12	6	1	–	–	–	–	41	333	2799
16	–	–	–	–	–	3	9	12	5	4	10	1	1	45	487	5413
17	–	1	2	2	6	5	10	2	3	–	–	–		31	251	2125
18	–	–	–	–	1	1	2	1	–	–	–	–	–	5	43	375
19	–	–	4	4	11	8	7	2	–	–	–	–	–	36	268	2062
20	–	–	–	5	2	10	9	14	8	2	–	–	–	50	457	4303
21	–	–	1	1	10	13	10	2	–	1	–	–	–	38	307	2537
22	–	–	1	3	7	15	13	7	4	–	–	–	–	50	423	3673
23	–	–	–	3	14	15	12	5	–	–	–	–	–	49	394	3226
24	–	–	–	1	8	14	12	9	2	1	–	–	–	47	406	3582
25	–	–	1	3	19	11	8	3	1	–	–	–	–	46	357	2837
26	–	–	1	1	2	1	11	16	12	3	–	1	–	48	474	4794
27	–	–	–	1	4	12	16	7	1	1	1	–	–	43	380	3430
28	–	–	–	1	4	4	16	10	2	–	–	–	–	37	332	3026
29	–	–	–	1	5	13	6	8	6	4	–	–	–	43	393	3701
30	–	–	–	1	5	9	14	9	6	–	–	–	–	44	395	3617
31	1	–	1	6	8	12	6	4	6	1	–	–	–	45	368	3166
32	–	–	–	–	6	10	9	7	1	–	1	–	–	34	297	2653
33	–	–	1	4	7	21	10	5	–	–	–	–	–	48	386	3166
34	–	–	–	1	9	15	11	7	1	–	–	–	–	44	369	3149
35	–	–	4	3	9	9	8	10	2	1	–	–	–	46	379	3259
36	–	–	–	1	2	14	9	9	6	–	–	–	–	41	369	3385
37	–	–	–	2	6	13	7	7	–	4	–	–	–	39	339	3041
38	–	–	1	2	10	13	14	2	2	1	–	–	–	45	371	3139
39	–	–	–	–	7	15	19	4	–	–	–	–	–	45	380	3242
40	–	–	–	–	2	10	10	7	8	–	–	–	–	37	342	3216

These formulae lend themselves readily to practical manipulation. The rough moments for the parental-offspring surface are generally available; the others are obtained by rapid machine calculation.

Illustration (V) 11. *Avuncular Correlation in Garden Beans.*

As an illustration I take the calculation of the "avuncular" correlation in a series of forty families of garden beans—The Navy H series of former papers*. Table XIV gives the necessary data. The first part shows the weight of the seeds from the 40 parents, seriated individually with the sums and sums of squares of their values. Only a portion of the seeds (roughly 10) from each of these mothers were planted, and give the offspring arrays shown with their summations in the second part of the table.

The direct parental correlation†—that between the weight of the individual seed planted and the weight of the seeds which it produces—is given in Table XV.

TABLE XV.

Seed Weight in Garden Beans.

Offspring Weight.

Parental Weight.	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	Totals
5	—	—	—	—	1	4	—	2	1	—	—	—	—	8
6	—	—	2	6	8	4	4	4	—	—	—	—	—	28
7	—	—	1	7	21	18	21	4	3	4	—	—	—	79
8	—	2	10	21	62	103	100	55	19	4	3	—	—	379
9	1	1	15	29	74	137	134	72	23	3	—	—	—	489
10	1	3	7	30	71	102	77	54	25	5	—	—	—	375
11	1	—	6	18	31	51	66	47	22	12	2	2	—	258
12	—	1	—	—	2	7	14	10	3	2	3	—	—	42
13	—	—	—	—	—	—	—	3	2	—	4	—	1	10
Totals	3	7	41	111	270	426	416	251	98	30	12	2	1	1668

* *American Naturalist*, Vol. XLVI. pp. 313—343, 512—525, 656—674, 1912; Roux's *Archiv f. Entwicklungsmechanik der Organismen*, Bd. XXXV. S. 500—522, 1912; *Biometrika*, Vol. IX. pp. 11—21.

† The Editor has called my attention to the possibility of some misunderstandings which might be introduced by this illustration and I am glad to add a word of explanation concerning it.

The chief question which is likely to be raised is: "What constitutes direct parental inheritance in a character like seed weight?" Some rather complicated biological questions are involved. The coats of the seed are a part of the plant individual which produces them—are in fact strict homotypes in Pearson's sense of the term. But the embryo—the cotyledons, hypocotyl and the plumule, *the elements which constitute the major part of the weight of the seed*—is strictly speaking a daughter plant. Hence an intra-individual correlation for seed weight is not a true homotypic correlation but a fraternal or sororal correlation. Thus from the embryological or genetic standpoint the correlation between the weight of the individual seed planted and the weights of the seed borne by the plant into which it develops is the true parental correlation.

There are, however, two important modifying conditions. First, the correlation is between two individuals in an embryonic stage, not in the adult condition. It is somewhat analogous to the correlation between the larval characteristics of two successive generations of moths, and resembles still more closely the coefficient measuring the resemblance between a parent's weight at birth and

Designating the parental generation seeds by x and the offspring seeds by y, I find from Table XV,

$$\Sigma (x') = 15416, \quad \Sigma (y') = 14097,$$
$$\Sigma (x'^2) = 145400, \quad \Sigma (y'^2) = 123273,$$
$$\Sigma (x'y') = 130794,$$

whence, with Sheppard's correction,

$$r = (130794/1668 - 9{\cdot}242206 \times 8{\cdot}451439)/1{\cdot}291727 \times 1{\cdot}547424 = {\cdot}1519.$$

From Table XIV I find for the ascendant-descendant correlation

$$S(pq) = 60756,$$
$$S[q\Sigma (x')] = 564066, \quad S[p\Sigma (y')] = 514990,$$
$$S[q\Sigma (x'^2)] = 5351080, \quad S[p\Sigma (y'^2)] = 4518074,$$
$$S[\Sigma (x') \Sigma (y')] = 4798171.$$

These give

$$r = (4798171/60756 - 9{\cdot}284120 \times 8{\cdot}476364)/1{\cdot}340412 \times 1{\cdot}559539 = {\cdot}1334.$$

But obtaining the summations about 0 for the avuncular relationship by simple subtraction I get for the population N

$$S[q(p-1)] = S(pq) - S(q) = 60756 - 1668 = 59088,$$

and for physical constants, for "aunts"

$$\bar{x} = (564066 - 15416)/59088 = 9{\cdot}285303,$$

$$\sigma_x = \sqrt{(5351080 - 145400)/59088 - (9{\cdot}285303)^2 - {\cdot}083333} = 1{\cdot}341743,$$

and for "nieces"

$$\bar{y} = (514990 - 14097)/59088 = 8{\cdot}477068,$$

$$\sigma_y = \sqrt{(4518074 - 123273)/59088 - (8{\cdot}477068)^2 - {\cdot}083333} = 1{\cdot}559872,$$

and for the relationship between them

$$r = [(4798171 - 130794)/59088 - \bar{x}\bar{y}]/\sigma_x \sigma_y = {\cdot}1329.$$

a child's weight at birth. Second, all the seeds produced by the same plant are under closely similar environmental conditions. Thus while they are (so far as the embryo to which weight is chiefly due) sister seeds, they are far more nearly alike than would be expected on the basis of heredity alone, for all are borne in and nourished by the maternal tissues.

Under these circumstances one may consider either (a) the correlation between the individual seed planted and the individual seed produced, or (b) the correlation between the average character of all (or of a representative sample) of the seeds of a parental and filial individual to be the true correlation.

If the latter view is taken one considers the characteristics of the seeds to be determined exclusively by the plant on which they are borne (the parent plant).

If the former alternative be adopted one neglects for the moment the important environmental factor introduced by the fact that the whole array of parental seeds are subjected to the limited environment of one plant while the offspring array (produced by any single plant) is subjected to the environment of another plant.

It has seemed to me safest to follow strictly the biological homologies in terminology.

I note that the parental relationship is the highest and the avuncular the lowest. These results cannot be discussed in detail until the final memoir from which they are abstracted is published, but it seems to me that they agree with Hanel's data for *Hydra* as analyzed by Pearson* in affording no support to Johannsen's conclusions concerning the ineffectiveness of selection within the pure line.

If one desires a check, or if linearity of regression must be tested, one may proceed by the method already described for the formation of condensed correlation tables †.

Either ascendant or descendant seeds may be taken as the "first" or subject character. Since they are quickly formed, both have been drawn up for the convenience of the reader and appear as Tables XVI—XVII. In them, the first two columns give the reduced tables for parents and offspring (or offspring and parents), the second two give the frequencies for the ascendants weighted with their descendants (or the descendants weighted with their ascendants) and the totals of the associated arrays. The final two columns are obtained by subtracting the first pair from the second, and give the avuncular relationship desired. Naturally, the final summations of these columns may be subtracted if linearity tests only are not desired. All the above constants may be checked directly from these tables.

TABLE XVI.

Weight of Ancestral Seed	Parental Correlation		Ascendant-Descendant Correlation		Avuncular Correlation	
	Frequency of Actual Parents Weighted with their Offspring	Total Offspring Weight	Weighted Frequency of Descendant Seeds	Total Weight of Weighted Descendant Seeds	Weighted Frequency of Nepotic Seeds	Total Weight of Weighted Nepotic Seeds
3	—	—	35	306	35	306
4	—	—	40	340	40	340
5	8	70	294	2526	286	2456
6	28	210	1147	9313	1119	9103
7	79	648	3389	27780	3310	27132
8	379	3188	11510	96824	11131	93636
9	489	4085	18095	151467	17606	147382
10	375	3111	15184	127618	14809	124507
11	258	2270	8247	71591	7989	69321
12	42	396	2406	22963	2364	22567
13	10	119	364	3775	354	3656
14	—	—	45	487	45	487
Totals	1668	14097	60756	514990	59088	500893

* *Biometrika*, Vol. VII. pp. 372—381, 1910.

† *American Naturalist*, Vol. XLVI. pp. 477—486, 1912.

TABLE XVII.

Weight of Offspring Seed	PARENTAL CORRELATION		ASCENDANT-DESCENDANT CORRELATION		AVUNCULAR CORRELATION	
	Actual Frequency	Total Weight of Weighted Parental Seeds	Weighted Frequency of Ascendant Seeds	Total Weight of Weighted Ascendant Seeds	Weighted Frequency of Avuncular Seeds	Total Weight of Weighted Avuncular Seeds
3	3	30	127	1213	124	1183
4	7	67	263	2514	256	2447
5	41	370	1473	13393	1432	13023
6	111	1012	3901	35562	3790	34550
7	270	2437	9672	88297	9402	85860
8	426	3892	15421	142308	14995	138416
9	416	3841	15190	139924	14774	136083
10	251	2366	9333	88150	9082	85784
11	98	939	3642	34725	3544	33786
12	30	293	1098	10830	1068	10537
13	12	134	511	5757	499	5623
14	2	22	82	891	80	869
15	1	13	43	502	42	489
Totals	1668	15416	60756	564066	59088	548650

III. RECAPITULATION.

The term intra-class is suggested for the correlation describing the degree of resemblance of the individuals of the series of sub-groups or sub-classes into which a general population may be divided for some logical (*e.g.*, biological, sociological) reason. The term inter-class is applied to those correlations which describe the degree of interdependence of variables belonging to the logically related sub-classes of different populations. Such coefficients have a very wide but hitherto little realized range of usefulness.

Generally, their determination requires the comparison of each individual with the array of individuals constituting its own or its associated class. Heretofore the use of such coefficients has been limited, indeed almost precluded, by the excessive labour of forming the weighted tables. Methods for the rapid formation of the tables by the aid of mechanical calculators have been described, and formulae are now given by which such correlations can be determined without the formation of tables at all.

The labour of calculation may thus be reduced to a small fraction of that formerly required. The formulae present numerous opportunities for checking the arithmetic, but of course certain errors may be more easily detected when tables are used. Furthermore the tables, full or condensed, should be formed in

a certain number of cases for the purpose of testing the form of the regression curve.

Finally, in emphasizing the use of Intra-class and Inter-class Coefficients it may not be out of place to warn the reader that, as in the case of many other statistical relationships, the *raw or uncorrected* values may give quite erroneous impressions of the biological relationships under investigation. Some such cases will be discussed later. In the meantime, the purpose of this paper will have been fulfilled if the range of illustrations proves suggestive of wider application and if the formulae prove as useful to other biologists as they have to workers in my laboratory, where they have been very extensively used and have become indispensable tools.

In conclusion I take great pleasure in thanking most heartily the Editor of this Journal for the kindly criticisms and suggestions which have materially improved the manuscript.

A SECOND STUDY OF THE WEIGHT, VARIABILITY AND CORRELATION OF THE HUMAN VISCERA.

BY M. GREENWOOD, JUNR., AND J. W. BROWN.

(*From the Statistical Laboratory of the Lister Institute.*)

IN 1904, one of us published an account of certain applications of biometric methods to data obtained from the pathological department of the London Hospital*. The object of that essay was to throw light upon the mean values, variabilities and correlations of the human visceral organs, in the hope of (*a*) Placing on record results comparable with the numerous studies of anthropologists which deal with other organs or dimensions: (*b*) Providing materials for the study of the changes produced by various diseases in the absolute measures and inter-relations of the organs observed and, therefore, *pro tanto*, illustrating the work of selection upon the human species.

In the paper cited, attention was directed to the specialised characters of a "General Hospital Population" and it was pointed out that results obtained from such a population could not be generalised without inquiry as applicable to the race at large. It was also shown that groups selected upon various principles within the "General Hospital Population" exhibited marked deviations from the prevalent type, a selection of the more healthy organs being attended by a rise in correlation and a fall in variation, while a group characterised by a preponderance of specifically diseased organs exhibited changes in the opposite sense.

The difficulty which had to be faced in that investigation was the task of ascertaining the characteristics of a healthy "General Hospital Population." By this expression we define the class of persons of the physical, racial and economic type to which hospital patients conform but who are not suffering from any disease likely to cause death or seriously affect health. The nature of the material available in 1904 did not permit of such a selection being made and all that could be done was to choose from the data those organs which, so far as the records showed, were not obviously diseased. Such cases formed the normal or "healthy" material

* *Biometrika*, Vol. III. p. 65.

of the earlier paper. In Table I we have grouped the results obtained from such data; the figures relating to males have already been published, those dealing with females are based upon the manuscript records, and are given for the sake of comparison. It will be noticed that in two cases the correlations for the female data

TABLE I.

A. *Constants calculated from Old Data. Healthy Organs.*

Organ	Sex	Mean	Standard Deviation	Correlation
Heart ...	M.	11·04 ± ·07 ozs.	1·923 ozs.	·2780 ± ·0329
Liver ...	M.	60·44 ± ·32 ,,	8·948 ,,	[H. and L.]
Heart ...	F.	10·12 ± ·11 ,,	3·087 ,,	·2163 ± ·0325
Liver ...	F.	55·23 ± ·15 ,,	4·298 ,,	[H. and L.]
Heart ...	M.	11·25 ± ·06 ,,	2·073 ,,	·2654 ± ·0276
Spleen ...	M.	5·22 ± ·06 ,,	1·996 ,,	[H. and S.]
Heart ...	F.	10·02 ± ·10 ,,	3·422 ,,	·1881 ± ·0295
Spleen ...	F.	6·07 ± ·10 ,,	3·197 ,,	[H. and S.]
Heart ...	M.	11·24 ± ·06 ,,	1·946 ,,	·4004 ± ·0279
Kidneys ...	M.	12·01 ± ·07 ,,	2·016 ,,	[H. and K.]
Heart ...	F.	9·45 ± ·09 ,,	2·445 ,,	·4465 ± ·0292
Kidneys ...	F.	10·58 ± ·10 ,,	2·620 ,,	[H. and K.]

B. *Relative Variability of Healthy Organs.*

Organs	Sex	Coefficient of Variation	
Hearts with Livers ...	M.	17·42	mean 17·71
Hearts with Spleens ...	,,	18·42	
Hearts with Kidneys ...	,,	17·30	
Livers	,,	14·80	
Spleens	,,	38·21	
Kidneys	,,	16·80	
Hearts with Livers ...	F.	30·49	mean 30·17
Hearts with Spleens ...	,,	34·17	
Hearts with Kidneys ...	,,	25·86	
Livers	,,	7·78	
Spleens	,,	52·69	
Kidneys	,,	24·76	

are slightly smaller, in one instance perhaps significantly smaller, than the male values, while the relative variabilities are in two cases higher among the females.

It was to be anticipated, from the nature of the case, that the results just given would be only rough approximations to the desired values descriptive of a healthy "General Hospital Population." This anticipation has been confirmed by other work. Thus Pearson has shown that the form of frequency distribution indicated by the graph of "healthy" hearts would suggest arbitrary limitation at each end of the distribution, some small and many large hearts having been unjustifiably

classed as abnormal*. It therefore seemed desirable to re-investigate this particular problem and a special motive for doing so now is that one of us hopes shortly to collaborate with Dr Mott in the reduction of similar but more extensive data collected in Claybury Asylum. To make this latter inquiry complete, some comparison of biometric constants derived from an insane population with those yielded by sane stock is essential. The sane population as a whole not being available, and the economic strata which provide general hospital and county asylum patients being similar, it seemed probable that an extension of the earlier inquiry would be of value. Under these circumstances recourse was once more had to the pathological department of the London Hospital. We took this step with the more confidence because we knew that since the date of the first investigation the pathological records of the London Hospital had been enormously improved, under the supervision of Dr H. M. Turnbull, Director of the Pathological Institute.

Dr Turnbull kindly supplied us with data prepared in accordance with his instructions and we desire to express our gratitude to him not only for placing the material at our disposal but for much friendly advice and criticism throughout the inquiry.

The data provided by Dr Turnbull include, with the exception of the weight of the spleen, all the measurements recorded in the former series and in addition the weight of the brain (in most cases), and the weight of the whole body and its length. We must, however, remark that the last mentioned datum, although included in one of our tables, has been shown to be unreliable, so that no sound conclusions can be based upon it. We shall refer to this point later.

With respect to the methods by which these data were prepared, we have somewhat exact details. Dr Turnbull permits us to quote the following passage from a private letter addressed to one of us.

"Before the bodies are put on the P.M. tables they are weighed and measured and the results are entered by the porters upon the tops of the sheets for the description of the necropsy. Each organ after it has been removed, cut and described is placed in the tray at the end of the table and left there until Dr Bartlett or I have gone through the case dictating the diagnosis of the case as a whole and of each organ. So soon as most of the organs have been completely separated the porters lift them from the trays, weigh them and write down the weights on the large slate in the P.M. room. Occasionally an assistant places an organ in the tray before slicing into it; it may therefore be weighed before it is sliced, but this is exceptional. The organs are almost always completely separated and therefore weighed before Dr Bartlett and I go through the case. In special circumstances they are placed in the trays without separation, e.g., in disease of the bile ducts the liver, pancreas, stomach and duodenum would be placed together in the tray without separation until Dr Bartlett or I examined them. Such organs would be

* *Biometrika*, Vol. IV. p. 183.

weighed later than the others. In cases of accident it would be extremely unlikely that such a proceeding would ever be necessary. You may take it then that in the cases sent to you the organs were weighed almost immediately after removal and examination by the assistant."

This material although naturally less extensive than that previously used is far more reliable. Dr Turnbull knew that we wished to study healthy organs and consequently the data received by us formed a selection not a random sample of the "General Hospital Population." We have, however, carried the process of selection yet further. We chose from the data all the males at ages 25—55 and rejected from this number (1) all that had died of a wasting disease, such as Cancer or Tuberculosis; (2) all cases in which a possibly pathological condition of the heart was noted, whatever the weight of the organ might be; (3) cases in which Syphilis might be suspected (those with record of aneurism in particular); (4) all cases with no recorded body-weight.

The cases without recorded body-weight have not been considered at all, but those rejected from the series on account of (1), (2) or (3) were subsequently added to the remainder and the constants of the whole series ultimately calculated. Our work therefore falls into two parts, the first being an analysis of some 78 doubly selected cases, the second a reduction of about 100 records selected also but not so stringently. An examination of the data which are published as an appendix to this memoir will enable the reader to appreciate our method; it may be remarked that certainly one and possibly two cases not admitted to the "double selection" should have figured there (one was left out by mistake), while some of those admitted, in particular the epileptics, ought possibly to have been omitted. On the whole, however, we think the "double selection" conforms pretty well to the type we had in mind. Needless to say, it falls far short of the ideal, since, not to mention other points, we have not been able to take account of occupation or race in view of the paucity of our material. It may also be mentioned that we have other data relating to persons outside the limits of age here adopted; the analysis of these cases will be dealt with in another communication.

The non-medical reader will perhaps wonder how it is that we think it worth while to deal with so few cases, and may object that the London Hospital is only one of a number of great medical charities and that the material from all of these should be pooled and treated as a whole. This would be so obviously the proper course were it practicable that we may be allowed to point out the difficulties in the way of its adoption. The effective absence of coordination between the medical schools, the diverse systems in force and the contempt of statistical data which up to recent times characterised all but a minority of the hospital staffs, would render any individual attempt to deal with combined records a troublesome and possibly futile undertaking. We are not speaking without knowledge of the statistical systems in vogue at a large hospital when we say that the present state of affairs is unsatisfactory. Every year tons of paper and gallons of ink are devoted to recording the experiences, medical, surgical and pathological, of the great teaching

institutions. An extremely small percentage of the results have any value at all, while even that residuum is not readily accessible. Comparatively little trouble and some expense would notably diminish the output of waste-paper and increase the production of valuable records.

Hospital administrators pay little attention to the advice of private enthusiasts —and quite properly since compliance with requests for innovation might turn hospitals into happy hunting grounds for faddists of all kinds. There are we think but two ways in which genuine statistical reforms are likely to be brought about. One is through Government or Municipal action, a process which does not (if we may judge from the report of the recent Royal Commission on the Record Office) necessarily lead to the best results. The second is through the realisation on the part of leading physicians and surgeons of the importance which attaches to organised statistical data. Perhaps when consumption has been "conquered" and the infant death rate reduced 80°/ₒ, medical protagonists may be induced to turn their attention to a humbler but still important reform. The question is really of more importance than might be inferred from the subject of this paper. The problem with which we are immediately concerned here is not of momentous interest, but the difficulties under which we labour apply to all statistical investigations of a "General Hospital Population." Since every youth destined to earn his bread as a doctor acquires his first and therefore his most vivid impressions of the phenomena of disease among a "General Hospital Population," it is of importance to all his future patients that these impressions should be controlled, checked and, if possible, corrected by any available method. One such method, the statistical realisation of precisely what a "General Hospital Population" is, and how it may be presumed to differ from the nation at large, is at present impossible owing, in great measure, to obstacles which can be removed. This is now the opinion of a small minority; if it is ever adopted by a majority, those who, having the means of reform at their command, neglect to adopt them will find some difficulty in explaining their position.

TABLE II.

Constants calculated from New Data (Select Cases). Healthy Organs. Males (25—55).

Variable	Mean	Standard Deviation	Coefficient of Variation
Heart	11·96 ± ·14 ozs.	1·896	15·86
Body-weight	129·82 ± 1·80 lbs.	23·605	18·18
Kidneys	10·67 ± ·16 ozs.	2·049	19·21
Liver	57·39 ± ·86 "	11·263	19·63
Age	41·27 ± ·66 yrs.	8·632	20·92
Brain	49·31 ± ·36 ozs.	4·558	9·24
["Body-length"] ...	[68·99 ± ·20 ins.]	[2·548]	[3·69]

We now pass to the results of our investigation and begin with the specially selected series, the chief constants of which are given in Tables II and III. All the mean values which can be compared with those set out in Table I differ significantly therefrom; the heart is heavier, the other organs lighter than in the first series. It is doubtful whether the changes in relative variability of the heart and kidneys are significant, the difference in the latter case being only 2·41 ± 1·12.

TABLE III A.

Total Correlations.

	Heart	Body-weight	Kidneys	Liver	Brain	["Body-length"]	Age
Heart ...	1	·65 ± ·04	·56 ± ·05	·52 ± ·06	·08 ± ·08	[·42 ± ·06]	·04 ± ·08
Body-weight ...	·65 ± ·04	1	·37 ± ·07	·57 ± ·05	·25 ± ·07	[·52 ± ·06]	− ·03 ± ·08
Kidneys ...	·56 ± ·05	·37 ± ·07	1	·43 ± ·06	·08 ± ·08	[·28 ± ·07]	− ·05 ± ·08
Liver ...	·52 ± ·06	·57 ± ·05	·43 ± ·06	1	·13 ± ·08	[·41 ± ·06]	− ·16 ± ·07
Brain ...	·08 ± ·08	·25 ± ·07	·08 ± ·08	·13 ± ·08	1	[·27 ± ·07]	− ·08 ± ·08
["Body-length"]	[·42 ± ·06]	[·52 ± ·06]	[·28 ± ·07]	[·41 ± ·06]	[·27 ± ·07]	[1]	[− ·20 ± ·07]
Age ...	·04 ± ·08	− ·03 ± ·08	− ·05 ± ·08	− ·16 ± ·07	− ·08 ± ·08	[− ·20 ± ·07]	1

TABLE III B.

Partial Correlations.

Variables	"Constants"	Correlation
Heart and Body-weight	Kidneys and Liver	·49 ± ·06
Heart and Kidneys	Body-weight and Liver	·41 ± ·06
Heart and Liver	Body-weight and Kidneys	·13 ± ·08
Body-weight and Kidneys	Heart and Liver	− ·07 ± ·08
Body-weight and Liver	Heart and Kidneys	·36 ± ·07
Kidneys and Liver	Heart and Body-weight	·21 ± ·07

The liver, on the other hand, seems to be definitely more variable in the present series. Our finding, so far as the heart is concerned, confirms the remark of Pearson already cited. With respect to the correlations, those which have corresponding entries in Table I are larger than these, the coefficient involving the liver being significantly increased. Some of the other constants, while not matched in Table I, may be compared with the reductions of Blakeman* and Pearl†, viz. brain- and body-weights. Our mean brain is heavier than that deduced by Blakeman from the Middlesex Hospital data and near Pearl's figure

* J. Blakeman, "A Study of the Biometric Constants of English Brain Weights," *Biometrika*, Vol. IV. p. 124, 1905.

† R. Pearl, "Variation and Correlation in Brain Weight," *ibid.* p. 13.

for Bohemian cases. The mean body-weight is greater than Pearl's figure and the coefficient of variation between that of Pearl and Pearson*.

Since Pearl and Blakeman did not restrict their inquiry to persons between 25 and 55 years of age, any comparison of our results with theirs must be of doubtful validity. Similarly we cannot control the constants involving age by comparisons with the work of other biometricians. In the first study, the correlation between age in years and "healthy" heart weight for males was given as $\cdot 136 \pm \cdot 025$ and from the same data we find in the case of females $r = \cdot 155 \pm \cdot 029$, in good agreement with the former value. Our present results cannot be said to differ significantly from the old values and we may quote a passage from the other memoir: "When we remember that the healthy heart is on the average much smaller than the heart in disease, and that sickness on the average increases continuously with age, we shall probably lay less stress on the general *a priori* idea that the weight of the adult heart increases very sensibly with age alone†." We see no reason to modify this view; further the recorded ages are likely to be subject to a considerable observational error in this class of material.

The general conclusion therefore both of the present and the first study is that age alone is not *proved* to be sensibly correlated with the weights of the viscera in an adult healthy "General Hospital Population."

The correlations involving body-weight are the most interesting in our new series since, with the exception of Pearl's results, they are so far as we are aware the first reductions of this variable in its association with the visceral weights which have been published. It may be freely admitted that the weight of a corpse must be so appreciably affected by the circumstances immediately preceding death that it cannot correspond at all closely to the weight during life. The force of this objection is diminished by the large percentage of violent deaths to be found in our series, since these cases would not be affected by the rapid wasting so often associated with a fatal illness.

Chiefly for this reason, we are not without hope that the present calculations throw some light upon the true correlation between heart weight and body-weight‡. One inference seems clear enough, viz. that no opinion ought to be ventured as to whether the weight of a heart is or is not normal without taking into account not only that of the other organs but also the body-weight. This leads to the question as to whether a formula of the nature of a regression equation might appropriately be employed to connect heart weight with other variables. In selecting variables upon which to base a prediction of the heart weight, the ideal is that each member of the set shall be highly correlated with the heart and that the correlation of the chosen variables one with another shall be low. For the purposes of a first trial, we decided to found the prediction

* K. Pearson, "The Chances of Death, etc.," Vol. I. pp. 293, etc.

† Greenwood, *op. cit.* p. 69.

‡ It is also to be remembered that there must be an appreciable *post mortem* loss of weight owing to evaporation in the cold chamber.

upon the weight of the whole body and that of the kidneys and liver. The required constants having been determined* we find:

$$x_1 = \cdot 0369x_2 + \cdot 3178x_3 + \cdot 0192x_4 + 2\cdot 6715 \quad \ldots\ldots\ldots\ldots\ldots\ldots(1),$$

where x_1 is the weight of the heart in ounces, x_2 the body-weight in pounds, x_3 the weight of the kidneys in ounces, x_4 the weight of the liver in ounces.

The standard deviation of the prediction is 1·235 ozs., giving a mean error of ·96 oz.

The value of such an equation depends upon the nature of the regression† which cannot be properly determined when the data are too few or too scattered to allow of grouping. A rough empirical test is to see how far the predicted value agrees with the known weight in a sample of the material. Blakeman applied this test in his series of brain weights. We took every third case in our selected series, determined the probable weight of the heart from the above equation and compared this with the recorded weight. The mean absolute error in the whole 26 was ·98 oz., which is in reasonable agreement with our expectation. We then asked ourselves whether the prediction could be improved.

In the Tables II and III we have recorded, within brackets and inverted commas, some constants of "Body Length." Inquiry has elicited the fact that the "body length" was not true height at all but the length from the crown of the head to the *end of the toes* and had been recorded for the purposes of the undertaker. Since this measurement will be greatly affected by the degree to which the toes are flexed when *rigor mortis* sets in it does not necessarily bear a constant proportion to the stature, as was readily seen on examination of a short series of cases upon which both measurements had been made. Consequently the mean and standard deviation of this measurement are of no interest; but, since we found it to be substantially correlated with other measurements, it seemed just possible that its inclusion in an equation would improve the prediction. We found:

$$x_1 = \cdot 0356x_2 + \cdot 3034x_3 + \cdot 0166x_4 + \cdot 0418x_5 + \cdot 2518 \quad \ldots\ldots\ldots\ldots(2),$$

where x_1 = heart weight in ounces, x_2 = body-weight in pounds, x_3 = kidneys in ounces, x_4 = liver in ounces, x_5 = "body length" in inches.

Applying this equation to the 26 test cases as before, the mean error proved to be ·97 oz., i.e. there was no sensible improvement in the fit.

Lastly we considered the effect of only taking two variables, the body-weight and the weight of the kidneys. The equation was:

$$x_1 = \cdot 0413x_2 + \cdot 3449x_3 + 2\cdot 9187 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3),$$

* In determining the various regression constants six places of decimals were retained in the coefficients of total correlation, i.e. we did not use simply the values tabulated in this paper. The work was arranged on the plan introduced by Yule, but several of the constants were checked by re-calculation, using the minors of the fundamental determinant in the ordinary way.

† An analysis of the data in Table VI of Greenwood's "First Study" shows that in the case of "healthy" organs the regression of heart weight upon kidney weight was effectively linear. For that case, $N=413$, $r=\cdot 4004$, $\eta=\cdot 4174$, and $\frac{\sqrt{N}}{\cdot 67449} \times \frac{1}{2}\sqrt{\eta^2 - r^2} = 1\cdot 776$.

where x_1 is the weight of the heart in ounces, x_2 the weight of the body in pounds, x_3 the weight of the kidneys in ounces.

(3) applied to the 26 cases gave a mean error of 1 ounce.

We accordingly conclude that the weight of the heart may be predicted from that of the body and kidneys with an average error of some 8 per cent. of the mean weight of the heart, and that the introduction of the liver does not appreciably improve the accuracy of the result. This equation may be of use occasionally in the *post mortem* room; rough as it is, we do not think any better formula has been found by the anatomist.

We finally come to the results yielded by the whole of the data sent us which fell between the required limits of age.

The necessary constants are set out in Tables IV and V A and B and these should be compared with Tables II and III. It will be seen at once that,

TABLE IV.

Constants calculated from All new Data. Males (25—55).

Variable	Mean	Standard Deviation	Coefficient of Variation
Heart	12·15 ± ·18 ozs.	2·700	22·22
Body-weight	128·89 ± 1·64 lbs.	24·367	18·91
Kidneys	10·77 ± ·15 ozs.	2·267	21·05
Liver	57·11 ± ·81 „	11·890	20·82
Age	42·60 ± ·60 yrs.	8·844	20·76
Brain	49·10 ± ·30 ozs.	4·464	9·09

TABLE V A.

Total Correlations. All new Data. Males (25—55).

	Heart	Body-weight	Kidney	Liver	Brain	Age
Heart ...	1	·56 ± ·05	·53 ± ·05	·46 ± ·05	·09 ± ·07	·06 ± ·07
Body-weight...	·56 ± ·05	1	·40 ± ·06	·57 ± ·05	·24 ± ·07	− ·04 ± ·07
Kidney ...	·53 ± ·05	·40 ± ·06	1	·32 ± ·06	·10 ± ·07	·04 ± ·07
Liver ...	·46 ± ·05	·57 ± ·05	·32 ± ·06	1	·14 ± ·07	− ·11 ± ·07
Brain ...	·09 ± ·07	·24 ± ·07	·10 ± ·07	·14 ± ·07	1	− ·08 ± ·07
Age ...	·06 ± ·07	− ·04 ± ·07	·04 ± ·07	− ·11 ± ·07	− ·08 ± ·07	1

TABLE V B.

Partial Correlations.

Variables	Constants	Correlation
Heart and Body-weight	Kidneys	·45 ± ·05
Heart and Kidneys	Body-weight	·40 ± ·06
Kidneys and Body-weight	Heart	·15 ± ·07

excepting the variability of the heart, no appreciable alteration can be discerned. Such changes as there are, e.g. the increased variability of the heart and its lessened correlation with the other organs, point to the first selection as being somewhat more healthy than the whole series, but the differences are evanescent. The regression equation of heart upon body and kidney weight as determined from the whole series is:

$$x_1 = \cdot 0462 x_2 + \cdot 4260 x_3 + 1 \cdot 6143 \quad \text{.........................(4)},$$

where x_1 = heart weight in ounces, x_2 = body-weight in pounds, x_3 = kidney weight in ounces.

The mean error for the 26 test cases is 1·1 ounces.

We think the inference to be drawn from this comparison is that, in order to arrive at the values of the true *post mortem* constants descriptive of a healthy "General Hospital Population," the process of selection need not be very stringent, so far as the diagnosed cause of death is concerned.

We also think the whole tendency of the present investigation is to suggest that the constants here published are, within the limits indicated by the errors of sampling, representative of a healthy "General Hospital Population." None of our results are in flagrant opposition either to the findings of observers equipped with more accurate data or to those previously obtained from hospital series. We hope, therefore, to have added another item of use to the student of a "General Hospital Population." Of course, being restricted to a single type of measurement, this study is of but limited interest. Shrubsall has published* the results of an inquiry into certain anthropological characters of a "General Hospital Population," while Greenwood and Candy† have shown cause for believing that the fatality of certain common diseases, such as lobar pneumonia, may be quite different in this class from that obtaining in the community at large. These are other items in a catalogue which has yet to be completed. It may well be possible for a future synthesist to paint a fairly accurate composite picture of a section of our countrymen which is of immense importance not only in itself but also as the substratum of medical instruction. At present, however, we must be contented with the task of collecting material, in particular statistical records.

The general conclusions of this inquiry may be summarised as follows:

(1) The mean weight of the heart in healthy adult males of the class from which patients in a London General Hospital are drawn is not far short of 12 ounces.

(2) The correlation between the weight of the heart and that of the whole body is not much less than ·5.

* F. C. Shrubsall, *Brit. Med. Journ.* Vol. II. p. 1689, 1904; *St Barts. Hosp. Reports*, Vol. XXXIX. p. 64.

† *Journ. Roy. Statist. Soc.* Vol. LXXIV. p. 365, 1911.

(3) The weight of the heart can be deduced from the weight of the body and that of the kidneys by means of a linear equation with an average error of about eight per cent.

(4) The weight of the body and that of the kidneys ought always to be considered before deciding whether a given heart is abnormally large or small.

(5) It is doubtful whether the weight of the heart is sensibly correlated with age in an adult population (of ages 25—55).

(6) The weight of the brain is less closely associated with the body weight and that of other viscera than is the heart weight.

APPENDIX OF MEASUREMENTS.

Data for the Study of Visceral and Organic Correlations in Healthy Male Adults (25—55).

Case Number	Age (years)	Body-weight (lbs.)	[" Body-length " (ins.)]	Heart-weight (ozs.)	Brain-weight (ozs.)	Kidneys (ozs)	Liver (ozs.)	Cause of death
42, '12	45	140	69	11·5	51·5	5·25	57·5	Fract. Skull
68, '12	45	155	67	14·75	52·5	14·5	70	„ „
80, '12	38	138	69	13·5	52	9	42·5	Abdom. Injury
114, '12	41	117	67	10·5	49	9·5	50·5	Fract. Pelvis
178, '12	48	132	67	14·75	50	12·5	57·5	Tetanus
190, '12	50	136	68	13·5	46	11·5	64	Ruptured Intestine
253, '12	52	126	68·5	10·5	52·5	10·75	47	Fract. Skull
20, '11	43	112	68	9·5	52·5	11·25	68·5	Bronch. Pneumonia
23, '11	34	157	71	11·5	49·5	9·5	74·5	Appendicectomy
37, '11	45	128	67	12	48	11·5	74	Traum. Rup. of Oesoph. Pneu- [mothorax
61, '11	25	95	67	10·5	49·5	10	48	Periton. Perf. Duoden. Ulcer
133, '11	41	133	69	11·75	47·75	12·75	65·5	Strang. Hernia Perit.
144, '11	34	106	66	10	—	9·5	57	Perf. Gastric Ulcer
169, '11	54	138	70	14·5	—	13·5	90	Fract. Spine
174, '11	53	145	67	12	—	9	48	Appendicectomy
281, '11	30	160	73	11	54·5	9	52·5	Accident
292, '11	31	116	76	14	—	14·5	71·5	Glioma of Brain
374, '11	38	148	72	15	39·5	16·5	68	Anaesthetic Death
554, '11	54	136	67·5	12·25	47	9·75	64	Accident
629, '11	35	119	70·5	11	50	8	63	Tetanus
913, '11	46	128	72	12	54	9	59	Fract. Skull
1197, '11	54	100	66	9·75	43	8·5	34	Fract. Pelvis
1234, '11	50	115	69	11·75	39·75	9·5	49·5	„ „
1236, '11	49	121	68·5	13	50·25	11	56·5	Injury to Head
56, '10	52	149	72	14·25	52	10·25	70	Fract. Base
96, '10	34	168	72	13·25	47	9	75·5	Fract. Ribs
99, '10	26	95	66	9	43	7	50	Corrosive Poison
221, '10	49	126	67·5	12·75	46·5	11	66·5	Fract. Base
251, '10	47	108	69	9·5	51·5	9	55·5	Fract. Ribs
284, '10	36	188	71	13·5	48·5	11·5	79·5	Fract. Base
308, '10	40	98	68	10	43	10·5	48·5	Tetanus
403, '10	35	123	69	13	53	12·5	62	Fract. Skull
501, '10	46	102	64	12	46·5	14·25	34	Fract. Pelvis
522, '10	46	121	69	10·5	53	8·5	46	Epilepsy
608, '10	33	170	70·5	13·5	45	12	60	Fract. Skull
672, '10	52	129	71	11·5	53·5	11	54·5	Fract. Base

Data for the Study of Visceral Correlations—(continued).

Case Number	Age (years)	Body-weight (lbs.)	["Body-length" (ins.)]	Heart-weight (ozs.)	Brain-weight (ozs.)	Kidneys (ozs.)	Liver (ozs.)	Cause of Death
1047, '10	40	114	69	9·5	48	11·75	70·25	Fract. Skull
1048, '10	46	132	67·5	13	42·25	11·75	57·5	Fract. Base
1054, '10	51	156	67	15·5	49·5	12·5	70	Fract. Ribs
1159, '10	46	184	73	12·5	59	12·5	60	Shot
1160, '10	32	217	76	19·5	59	13·25	74·5	„
1172, '10	50	177	70	11·5	48	12·75	56·5	Fract. Ribs
1190, '10	49	112	66	9	44	9·5	46·5	Accident
226, '09	33	128	—	11·5	44·5	10·75	41	Fract. Base
294, '09	52	112	71	9·25	53	11	45	Fall
297, '09	45	114	69	9·75	48·5	9·5	42	Accident
298, '09	50	133	67	9·5	49	7	47	„
299, '09	50	127	71	14·25	47	13	66	„
329, '09	52	105	70	11·75	53	9	40	Oxalic Acid Poison
336, '09	29	121	69	10·75	54·5	12	55·5	Epileptic Fit
356, '09	48	108	69	11·5	49	9·5	50·5	Burns
365, '09	30	115	70	11·5	49	14	60	Tetanus
383, '09	26	135	73	14·5	44	9·75	53·5	Stabbed
449, '09	33	136	74·5	14·25	47·5	11	59	Tetanus
460, '09	43	122	72	11·5	50	11·25	64	Tetanus and Bronch. Pneumonia
865, '09	49	105	69	10·25	43	8	36	Cut Throat
1079, '09	48	127	66	13·5	56	11·5	53·5	Run Over
1261, '09	40	149	69	13	48·5	11	62·5	Accident
109, '08	26	105	65	10·5	46	10·5	62·5	Epileptic Fit [Peritonitis
145, '08	37	145	70	11·5	47	12	78	Anaesthetic Death, Appendic.
151, '08	43	96	65	9·5	47	8	43	Epileptic Fits
169, '08	35	105	64	12	44	8	50	Fract. Skull
172, '08	41	126	70	14·5	48	11·5	60	„ „
186, '08	25	130	70	11·5	52·5	12	60	„ „
202, '08	34	108	66	9	48	6·5	42	Scalds
207, '08	41	107	68	11	62·5	10	52	Fract. Skull
218, '08	44	147	68	10	51·5	9·75	63	Periton. Rup. Duod. Ulcer
221, '08	49	161	68·5	14·75	45·5	13·75	60	Fract. Base
384, '08	45	144	66	13·5	46	11	69	Burns
404, '08	32	136	65	12	52	13	—	Caught in Machinery
472, '08	26	148	73	11	55·5	10	57·5	Accident
505, '08	35	146	70	13	58·5	10	60	Opium Poisoning
542, '08	26	107	67	10·5	45	8·5	44·5	Fract. Skull
573, '08	35	169	71	14	—	14	73	Pulm. Emb. after Operation
680, '08	46	118	67	11	43	8·5	44	Fract. Spine
709, '08	26	122	67	10·5	55	12	49·5	Perf. Duod. Ulcer
751, '08	50	110	65	12	53·5	10	53	Tetanus
1014, '08	50	109	70·5	11·25	46	9	51·5	Fract. Spine
1104, '08	32	116	70	11	55	12·5	58·5	Stovain Poison

Classification of Causes of Death in above Data.

Accidents, Suicide and Homicide	49*
Intestinal Injuries and Post-Operative Effects	13
Tetanus	7
Poisoning	4
Epilepsy	4
Glioma of Brain	1
Bronch. Pneumonia	1
	79

* Fract. Skull 19. Other Fractures 11. "Accidents" 11. Homicide 4. Injury to Head 1. Burns and Scalds 3.

Other cases within the same limits of Age but excluded from the first Analysis.

Case Number	Age (years)	Body-weight (lbs.)	["Body-length" (ins.)]	Heart-weight (ozs.)	Brain-weight (ozs.)	Kidneys (ozs.)	Liver (ozs.)	Cause of Death
97, '09	50	166	74	12·75	48	14·25	92	Ruptured Aneurism
1244, '09	52	123	71·5	13	55	11	62·5	Ruptd. Gast. Art., Chr. Gast. Ulcers
1276, '09	28	108	71	8·5	48	9	45	Ruptd. Cong. Aneur. of Cereb. Artery
859, '09	54	128	74	13	—	9	56	Anaemia and Scars of Gast. Ulcers
810, '10	55	154	65	15*	48	13	59·5	Fract. Base (? Hyperpiesis)
15, '11	55	107	69	10·5	48	15·5	45·5	Operation. Cirrhosis of Liver
16, '11	53	191	72	12†	51	11·75	63	Operation. Enterectomy. Peritonitis
534, '11	52	102	65	11·75‡	45	8	43	Fits. Anaemia
73, '11	53	93	66	10	—	8·5	45	Carcinoma of Oesophagus
99, '11	45	174	71·5	28§	—	19·5	56·5	Heart Failure; Hyperpiesis
104, '11	53	105	71	10·25	54·5	10	41·5	Rupt. of Cereb. Aneurism
136, '11	41	104	69	10	—	10·25	55	B. Pna. Cerebral Tumour
600, '08	52	105	68	8	39·5	7	45	Rupt. of Aneurism
345, '11	46	137	69	17·5‖	47	11·5	60	Accident
42, '11	50	80	70	9·5	—	9	50	Carcinoma of Stomach
19, '11	43	129	72	21·5	—	8·75	94	Chron. Nephritis
30, '11	50	103	70·5	9·5	48	14·5	43	Diss. Tuberculosis
1204, '11	46	128	69	14¶	45·5	11	47·5	Cerebral Haemorrhage
1245, '11	53	126	72	13	—	12	52	Rupt. of Aortic Aneurism
651, '11	40	135	69·5	12·5	49	8·75	57·5	? Hydrocyanic Poison; ? Epilepsy
77, '12	27	129	69·5	10·25	48	9·25	64	Fract. Base

* Slight Hypertrophy. † Brown Atrophy. ‡ Fatty Degeneration. § Hyperpiesis. ‖ ? Hyperpiesis. ¶ Hyperpiesis.

THE MEASUREMENTS OF THE PELVIS, WITH SPECIAL REFERENCE TO OBSTETRIC PREDICTION.

By D. H. DE SOUZA, Assistant Physician to the Westminster Hospital.

CONTENTS.

(1) *Introductory.*

The publication by Emmons* of a number of measurements of the female pelvis, carefully taken by himself, has afforded us an opportunity of investigating, by modern statistical methods, the relationship of the diameters to one another and to the bone measurements. As far as we can ascertain this has not previously been attempted. The subject has interested us not only from the statistical and anthropological sides but especially on account of its importance in practical obstetrics. Some of the diameters of the female pelvis can easily be measured on the living subject, others not at all, or only with great inconvenience to the patient. Unfortunately the two very important diameters at the pelvic brim, the obstetric conjugate and the transverse diameter, fall within the latter category. Empirical rules have been given, for example, to determine the obstetric conjugate from the diagonal conjugate, a diameter comparatively easily measured, but these lack both mathematical treatment and adequate statistical support. We have, in the following paper, considered methods of calculating these two important diameters from other pelvic measurements more readily obtainable. After making due allowance for moisture and soft parts we have left our results in a form which can be applied to the measurements as taken from the living subject.

* *Biometrika*, Vol. IX. p. 34, 1913.

Material. Emmons measured 217 pelves of the American Indian squaw in different museums in the United States of America. Our calculations have been made from these measurements; 216 pelves were considered in most cases, No. 115 in Emmons's Table being omitted as the diagonal conjugate was not stated. In calculations involving the pubic height only 212 were available, as this measurement was not given in Nos. 13, 72, 79, and 132 in his Table.

For the definitions of the different diameters and the way in which each was measured reference must be made to Emmons's paper. The names used by him have been retained with one exception. The term *oblique diameter*, as it has more than one significance, has been replaced by *diagonal conjugate*. The latter name is that usually given in this country to the diameter in question (sacro-subpubic) and is less likely to cause confusion.

Two modern English pelves, one male and one female, and two ancient Egyptian female pelves, kindly lent by Prof. Thane and Dr Derry respectively, were measured in the dry and moist state to ascertain differences in the diameters under these conditions. The measurements of these pelves and of an ancient Egyptian male pelvis, the property of Prof. Pearson, were employed to test the formulae calculated from the American Indian pelves.

For the numbers showing the thickness of the soft parts over the hip bone we are indebted to Dr Derry. They were obtained from measurements taken on 20 post mortem subjects. The maximum inter-cristal diameter was measured with callipers. As will appear later, a knowledge of the height of the hip bone may be of service in calculating diameters. This height, however, would vary in the living subject according to the position of the pelvis. If a woman be seated on a table it is easy to measure the maximum height of the iliac crest above the table. This could readily be done by the obstetrician if the measurement proved useful to him. It was therefore decided to consider this a measure of the height of the hip bone in the moist condition and covered with tissues. With the thigh of the corpse flexed to the sitting posture and the callipers pressed from above on to the iliac crest and from below on to the ischial tuberosity, this height is obtained. Dr Derry soon found, however, that he got the same result if he took the measurement with the subject supine, one limb of the callipers being pressed down from above on to the iliac crest as before, the instrument crossing in front of the groin, and the other limb passing between the thighs to reach the ischial tuberosity from below. Dr Braxton Hicks has confirmed this independently on three subjects. Accordingly most of the measurements have been taken in this manner.

The thickness of the soft parts was measured by means of an instrument with a rounded blade about as thick as a coarse darning needle but with not so fine a point. This was fitted into a hollow handle into which it could be pushed against a spring. A graduated scale showed the length of needle protruding from the handle. When in use the blade was passed through the skin and soft parts till its point was in contact with the bone. The handle was then pushed firmly

against the skin so that the only part of the blade protruding from it was that extending from the skin to the bone. The length of this was read off and gave the thickness of the soft parts covering the bone. This was measured over the ischial tuberosity, and over the iliac crest in vertical and horizontal directions. The first two measurements, subtracted from the measurement already obtained for the height of the hip bone covered with tissues, will give the height of that bone in the moist state. By subtracting the thickness of tissues over the ilium in a horizontal direction from the maximum inter-cristal diameter previously measured, the length of that diameter in the moist pelvis free from tissues can be ascertained.

The instrument was designed and made for us by Mr H. E. Soper of the Biometric Laboratory.

(2) *The relationship of the diameters and bone measurements to one another.*

The relationship of the diameters to one another was studied by means of the coefficient of correlation. As a preliminary, the average length and variability, as measured by the standard deviation, were found for each one. Emmons has already given the averages. Our results differ from his in some cases, probably owing to the fact that we divided the data into a larger number of groups. The diameters and bone measurements considered, together with the average and standard deviation of each, are shown in the following table. For the purpose of comparing the variabilities, the coefficients of variation are also given. All measurements in this and subsequent tables are expressed in centimetres.

TABLE I.

	Average	Standard Deviation	Coefficient of Variation
Inter-cristal	25·77 ± ·07	1·44 ± ·05	5·57 ± ·18
Inter-spinous	22·66 ± ·08	1·65 ± ·05	7·28 ± ·24
Transverse (brim)	12·97 ± ·03	·72 ± ·02	5·56 ± ·18
Diagonal conjugate ...	12·14 ± ·05	1·02 ± ·03	8·44 ± ·28
Obstetric conjugate ...	10·32 ± ·05	1·10 ± ·04	10·62 ± ·35
Antero-posterior (outlet) ...	11·61 ± ·04	·90 ± ·03	7·76 ± ·25
Inter-tuberal	9·78 ± ·04	·90 ± ·03	9·22 ± ·30
Posterior sagittal	7·56 ± ·04	·82 ± ·03	10·89 ± ·36
Pubic height	2·81 ± ·02	·52 ± ·02	18·30 ± ·62
Right hip—height	19·26 ± ·05	1·02 ± ·03	5·27 ± ·17
„ „ breadth ...	14·52 ± ·04	·87 ± ·03	5·97 ± ·19
Left hip—height	19·32 ± ·05	1·03 ± ·03	5·32 ± ·17
„ „ breadth	14·49 ± ·04	·83 ± ·03	5·72 ± ·19
Sacrum—height	10·00 ± ·05	1·15 ± ·04	11·52 ± ·38
„ breadth	11·47 ± ·03	·66 ± ·02	5·71 ± ·19

The pubic height is the most variable of all the measurements. Of the diameters the transverse and inter-cristal show the least variability. The height

and length of the hip bone and the breadth of the sacrum have coefficients not very different from those of the transverse and inter-cristal diameters, and the height of the sacrum has a variability just about double that of any other bone measurement. The results are not surprising. Apart from the variation in a measurement due to individual differences, there are other causes for variation in the case of these measurements taken on old pelves. It is not always easy to obtain accurate measurements, especially of the diameters, and after death parts of the bone get worn away. The wearing away of the pubic crest, for example, and the difficulty which may be experienced in determining the lowest point of the symphysis, must be accountable to a great extent for the large variability of the pubic height. The height of the sacrum, being dependent on the number of vertebrae united to form that bone and on the integrity of the tip, which is apt to get worn, is more variable than the breadth of that bone and than the height and breadth of the hip bone. The extremities of these last are rounded and less likely to be worn. The breadth is a little more variable than the height, and experience proves that, of the two, it is the more difficult to measure accurately.

The diameters must depend on the same factors as the bones, the individual differences, wear and tear, and difficulty of getting accurate measurements. The last especially must be of importance here. The obstetric conjugate, for example, depends upon the position of the sacral promontory, and therefore on the number of vertebrae forming the sacrum and on the proper fitting together of the bones, on the integrity of the upper and posterior part of the pubic bone, and on the ease or difficulty with which its anterior extremity can be determined. How much of the variation is due to each of these factors it is impossible to say. In such a diameter as the transverse there is little likelihood that the parts of the bone forming its extremities will be much worn away, but it would be unwise to go beyond a general statement of this kind.

The pubic height and the height of the sacrum vary more than any other measurement of a single bone with which we are acquainted. The measurements of the mandible come next in order. Their coefficients of variation, as given by C. D. Fawcett*, are, for the greatest height 9·93, for the greatest width at the condyles 7·46, and for the greatest width at the angles 7·62. The variations in the lengths of the long bones in different races (found for the French and Aino by Lee and Pearson†, for the Naqada by Warren‡, and for other races by Pearson§) have been compared with those of our measurements. The coefficients have as their limits 4·17 and 7·00, most of them being less than 5·5; hence the measurements of the pelvic bones, other than the pubic height and the height of the sacrum already considered, although in some few cases varying to about the same extent as the lengths of the long bones, have, on the whole, a tendency to be more variable.

* *Biometrika*, Vol. I. p. 408, 1902.

† *Proc. Roy. Soc.* LXI. p. 343, 1897.

‡ *Phil. Trans. Roy. Soc.* Series B. CLXXXIX. p. 135, 1897.

§ *The Chances of Death*, Vol. I. p. 256. London, 1897.

The coefficients of variation of the diameters were compared with those of the other recorded measurements which include more than one bone. The coefficients of the greater number of the skull measurements* lie between 3 and 5. There are a few which are more nearly equal to those in Table I. Some of the measurements of the face, for example, have coefficients lying between 6 and 8, and that of the breadth of the palate is 9·29. The diameters of the pelvis are, however, more variable than the bulk of the cranial measurements.

The average bone measurements given in Table I are also of interest. The asymmetry of the female pelvis has been pointed out from time to time by anatomists and has been strongly insisted upon by Hasse†. According to him the height of the left side of the pelvis is greater than that of the right, while the breadth is less. The numbers in the table bear out this statement, but, while Hasse makes these differences about 1 cm. each, the table shows them to be ·06 and ·03 cm. respectively. In this connection it may be noted that, in the 20 subjects measured by Dr Derry, the average height of the left hip bone in the moist condition was ·15 cm. greater than that of the right. Of the three dry female pelves used by us for testing purposes one followed Hasse's rule but the others varied, so that the average of the three showed the height of the right hip bone to be ·05 and its breadth ·04 cm. greater than that of the left. The number of negative results, though worthy of record, is, however, too small to place against the positive findings.

The coefficients of correlation were next found for every diameter taken with every other diameter in turn. The results are summarised in the table.

TABLE II.

	Inter-crests	Inter-spines	Trans-verse	Diagonal conjugate	Obstetric conjugate	Antero-posterior	Inter-tubers	Post-sagittal	Pubic height
Inter-crests	1	·79 ± ·02	·61 ± ·03	·24 ± ·04	·17 ± ·04	·30 ± ·04	·10 ± ·05	·11 ± ·05	·28 ± ·04
Inter-spines	·79 ± ·02	1	·52 ± ·03	·16 ± ·04	·13 ± ·05	·18 ± ·04	·07 ± ·05	·13 ± ·05	·22 ± ·04
Transverse	·61 ± ·03	·52 ± ·03	1	·09 ± ·05	·07 ± ·05	·35 ± ·04	·32 ± ·04	·17 ± ·04	·10 ± ·05
Diagonal conjugate ...	·24 ± ·04	·16 ± ·04	·09 ± ·05	1	·91 ± ·01	·31 ± ·04	·20 ± ·04	·21 ± ·04	·14 ± ·05
Obstetric conjugate ...	·17 ± ·04	·13 ± ·05	·07 ± ·05	·91 ± ·01	1	·30 ± ·04	·19 ± ·04	·21 ± ·04	·05 ± ·05
Antero-posterior ...	·30 ± ·04	·18 ± ·04	·35 ± ·04	·31 ± ·04	·30 ± ·04	1	·32 ± ·04	·49 ± ·03	·07 ± ·05
Inter-tubers ...	·10 ± ·05	·07 ± ·05	·32 ± ·04	·20 ± ·04	·19 ± ·04	·32 ± ·04	1	·36 ± ·04	− ·06 ± ·05
Post-sagittal	·11 ± ·05	·13 ± ·05	·17 ± ·04	·21 ± ·04	·21 ± ·04	·49 ± ·03	·36 ± ·04	1	·004 ± ·05
Pubic height ...	·28 ± ·04	·22 ± ·04	·10 ± ·05	·14 ± ·05	·05 ± ·05	·07 ± ·05	− ·06 ± ·05	·004 ± ·05	1

* Fawcett, C. D., *loc. cit.*

† *Archiv für Anat.* p. 244, 1891. *Ibid.* p. 1, 1910.

As the obstetric conjugate and the transverse diameter at the brim have to be obtained in practice not by direct measurement but by calculation, and as a knowledge of them is important and often necessary, special attention should be paid to these two diameters. The highest correlation (·91) exists between the obstetric and diagonal conjugates, and this justifies the reliance always placed by the obstetrician on the use of the latter to determine the former. The obstetric conjugate shows a moderate degree of correlation (·30) with the antero-posterior diameter at the outlet, but the coefficient is small between it and each of the other diameters, being very low for the transverse diameter (·07) and pubic height (·05). The diameters with which the transverse diameter at the brim is most highly correlated are the inter-cristal (·61) and inter-spinous (·52). Next in order are the antero-posterior at the outlet (·35) and the inter-tuberal (·32). With the other diameters the correlation is low.

In view of the fact that the pubic height forms one side of a triangle, of which the obstetric and diagonal conjugates form the other two sides, it is sometimes taken into consideration in calculating the obstetric conjugate from the diagonal conjugate. Thus, according to Waldeyer*, "in order to obtain the obstetric conjugate from the diagonal conjugate 1·75 to 2 cm., on an average, must be subtracted, and, the more acute the angle between the diagonal conjugate and the symphysis, and the higher the symphysis, the more must be taken away." It will be seen, on referring to the table, that the pubic height is not highly correlated with any diameter, and that with the obstetric and diagonal conjugates the coefficients are ·05 and ·14 respectively.

In order to arrive at formulae for finding the diameters it is necessary to use the measurements with which they are most closely correlated. The coefficients of correlation of the obstetric conjugate and the transverse diameter at the brim with the bone measurements were therefore found, to see whether any of them would be high enough to be of service. These are now given :

		Obstetric conjugate	Transverse
Right hip—height	...	$\cdot55 \pm \cdot03$	$\cdot48 \pm \cdot04$
" " breadth	...	$\cdot48 \pm \cdot04$	$\cdot45 \pm \cdot04$
Left hip—height	...	$\cdot56 \pm \cdot03$	$\cdot45 \pm \cdot04$
" " breadth	...	$\cdot51 \pm \cdot03$	$\cdot48 \pm \cdot04$
Sacrum—height	...	$\cdot31 \pm \cdot04$	$\cdot21 \pm \cdot04$
" breadth	...	$\cdot37 \pm \cdot04$	$\cdot56 \pm \cdot03$

All these measurements are more highly correlated with the obstetric conjugate than is any diameter except the diagonal conjugate (·91). With the transverse diameter the measurements of the hip bone show higher correlation than do any diameters except the inter-cristal (·61) and inter-spinous (·52), and the coefficient of correlation of the breadth of the sacrum with this diameter (·56) is exceeded only by that of the inter-cristal with it. Of these measurements, however, the only one which can easily be made with sufficient accuracy on the living subject

* *Das Becken*, p. 49. Bonn, 1899.

is the height of the hip bone. This can be taken in the sitting posture as already indicated. Since it is one of the measurements most highly correlated with both the diameters under consideration, it will have to be taken into account in working out the formulae from which these diameters are to be found. Its coefficients of correlation with the diameters most closely correlated with the obstetric conjugate (diagonal conjugate) and the transverse diameter (inter-cristal and inter-spinous) will therefore be required. These have been calculated for both hip bones.

	Right hip—height	Left hip—height
Diagonal conjugate	·63 ± ·03	·64 ± ·03
Inter-crests	·56 ± ·03	·57 ± ·03
Inter-spines	·44 ± ·04	·46 ± ·04

Apart from the closeness of the figures on the two sides these results call for no particular notice.

As a matter of anthropological interest the relationship of the bone measurements to one another was studied, and a table of the coefficients of correlation drawn up on a plan similar to that employed for the diameters. The correlation is greater between corresponding measurements of the hip bones on the two sides than between the height and length of the same bone. Thus the heights of the two hip bones have a coefficient ·98 and the breadths a coefficient ·94, while that for the height and breadth of the right hip bone is ·78 and for the same measurements on the left side ·81. This closer relationship between corresponding measurements on the two sides than between measurements on the same side has been noted by previous observers for other bones*. The crossed correlation between the height of one bone and the breadth of the other is not very different from the direct correlation between the height and breadth of the same bone, e.g. height of right hip bone with breadth of left bone gives a coefficient ·79, height of right bone with breadth of right bone a coefficient ·78, height of left

TABLE III.

	Right hip—height	Right hip—breadth	Left hip—height	Left hip—breadth	Sacrum—height	Sacrum—breadth
Right hip—height ...	1	·78 ± ·02	·98 ± ·002	·79 ± ·02	·34 ± ·04	·59 ± ·03
Right hip—breadth ...	·78 ± ·02	1	·77 ± ·02	·94 ± ·005	·31 ± ·04	·55 ± ·03
Left hip—height ...	·98 ± ·002	·77 ± ·02	1	·81 ± ·02	·37 ± ·04	·61 ± ·03
Left hip—breadth ...	·79 ± ·02	·94 ± ·005	·81 ± ·02	1	·30 ± ·04	·58 ± ·03
Sacrum—height ...	·34 ± ·04	·31 ± ·04	·37 ± ·04	·30 ± ·04	1	·32 ± ·04
Sacrum—breadth ...	·59 ± ·03	·55 ± ·03	·61 ± ·03	·58 ± ·03	·32 ± ·04	1

* Lewenz and Whiteley, *Biometrika*, Vol. I. p. 345, 1902.

bone with breadth of right bone a coefficient ·77, and height of left bone with breadth of left bone a coefficient ·81. Note also that there is no great difference between the coefficients of the corresponding measurements of the right and left hip bones with the sacrum (see first four numbers in the last row or column in the table).

Of only one pelvic bone, the sacrum, have the coefficients of correlation of the measurements been recorded previously. Warren*, studying the bones of the Naqada race, found that the coefficient of correlation for the height and breadth of this bone was ·46 for 32 males examined and ·31 for 45 females. The latter number is practically identical with our finding (·32). The coefficients given by him to show the correlation of the measurements of the long bones with one another lie between ·82 and ·98 for the male, and ·70 and ·96 for the female, with the exception of that for the maximum length of the clavicle with the oblique length of the humerus, which is ·68 for the male and ·53 for the female. Lewenz and Whiteley†, for the bones of the hand, found coefficients with a range ·43 to ·90 for measurements of the same finger, and ·59 to ·95 for other measurements on the same hand. If we omit coefficients between corresponding measurements on the two sides our coefficients have a range ·30 to ·79, and, on the whole, are lower than those between the measurements of the long bones together and of the bones of the hand together, being nearer to the latter than to the former. They are, however, greater than those between the diameters of the skull which C. D. Fawcett‡ found to lie between ·27 and ·49 for the male and ·12 and ·28 for the female.

It is interesting to note here that the coefficient between the length and breadth of the skull is ·34 in the male and ·14 in the female, while that between the length and breadth of the pelvis at the brim, i.e. between the transverse diameter and the obstetric conjugate, is ·07, and at the outlet, i.e. between the antero-posterior and inter-tuberal diameters, ·32.

(3) *The finding of the obstetric conjugate and the transverse diameter for the dry pelvis.*

The method adopted to determine the non-measurable from the measurable diameters was that employed by Prof. Pearson§ for obtaining the stature of an individual from the lengths of the long bones. The equation of the regression line was used for this purpose. In the equations which follow the lengths of the diameters are represented by letters as shown below:

Obstetric conjugate	C	Inter-tubers	I
Transverse	T	Inter-crests	K
Diagonal conjugate	D	Inter-spines	S
Antero-posterior	A	Pubic height	P

Height of right hip R
" " left " L

* *loc. cit.* † *loc. cit.* ‡ *loc. cit.*

§ *Phil. Trans. Roy. Soc.* Series A. Vol. CXCII. p. 169, 1898.

(A) *Obstetric conjugate.* Equations were obtained for deriving this from each of the diameters with which it is most closely correlated, viz. the diagonal conjugate (·91), and the antero-posterior (·30). These equations are

$$C = \cdot 978 D - 1\cdot 553 \pm \cdot 299 \quad \ldots\ldots\ldots\ldots (a),$$

$$C = \cdot 365 A + 6\cdot 081 \pm \cdot 705 \quad \ldots\ldots\ldots\ldots (b),$$

the first of which has a much smaller probable error (·299) than the second (·705). A single equation was also obtained, using both diameters

$$C = \cdot 972 D + \cdot 023 A - 1\cdot 743 \pm \cdot 298 \quad \ldots\ldots\ldots\ldots (c)$$

with an extremely slight improvement of the probable error as compared with (*a*).

The diameters next in order of correlation with the obstetric conjugate are the posterior sagittal (·21) and the inter-tuberal (·19). The posterior sagittal is not measurable on the living subject. The inter-tuberal was measured in a special way on the dry pelvis by Emmons and is not necessarily the same as the inter-tuberal diameter as understood by the obstetrician. The method is not applicable to the living subject, but this diameter was employed here to test whether it was of sufficient importance for its measurement to be followed up further. An equation was formed by taking it with the diagonal conjugate and the antero-posterior diameter

$$C = \cdot 972 D + \cdot 023 A + \cdot 001 I - 1\cdot 745 \pm \cdot 299 \quad \ldots\ldots\ldots\ldots (d).$$

On comparing (*c*) with (*d*) it is apparent that the addition of the inter-tuberal diameter makes little difference, and that little to the bad, as is shown by the probable errors. This diameter will therefore not be considered further.

It has been pointed out previously that the pubic height is sometimes used in calculating the obstetric conjugate but that its correlation with this diameter is low (·05). The equation connecting the obstetric conjugate with the pubic height is

$$C = \cdot 110 P + 10\cdot 011 \pm \cdot 738 \quad \ldots\ldots\ldots\ldots (e),$$

having the largest probable error (·738) of the equations yet obtained. The result of combining the pubic height with the diagonal conjugate is, however, good:

$$C = \cdot 989 D - \cdot 165 P - 1\cdot 229 \pm \cdot 293 \quad \ldots\ldots\ldots\ldots (f)$$

with a probable error lower than that of (*c*). This is due to the fact that, although the obstetric conjugate is more highly correlated with the antero-posterior diameter (·30) than with the pubic height (·05), the former of these two is more highly correlated with the diagonal conjugate (·31) than is the latter (·14), and it has been found that when the crossed correlation between the variables is high the resulting equation is not so good as when the crossed correlation is low. The pubic height, whether measured externally or internally, involves exposure of or inconvenience to the patient, so that it is desirable to avoid its use if just as good or nearly as good results can be obtained by other means.

Of the measurements which can be taken on the living subject, next to the diagonal conjugate the height of the hip bone shows the closest correlation with the obstetric conjugate (right ·55, left ·56). Equations were formed using the height of each hip bone

$$C = \cdot590R - 1\cdot037 \pm \cdot619 \quad \text{.........................}(g),$$

$$C = \cdot602L - 1\cdot310 \pm \cdot610 \quad \text{.........................}(h).$$

The equations as tested by probable errors are better than (*b*) and (*e*) but not as good as (*a*), (*c*), (*d*), and (*f*), but if each of the measurements be taken with the diagonal conjugate the resulting equations are exceedingly good

$$C = 1\cdot008D - \cdot049R - \cdot977 \pm \cdot298 \text{.........................}(i),$$

$$C = 1\cdot005D - \cdot042L - 1\cdot075 \pm \cdot298 \quad \text{.....................}(j).$$

The probable errors to four places would be ·2976 and ·2978 respectively while that of (*c*) is ·2984, so that the probable errors are less than that of any equation except (*f*). These equations have a practical advantage over (*f*), for the height of the hip bone can be easily measured on the living subject whilst, as already mentioned, there are disadvantages attached to the measurement of the pubic height. Moreover, the measurement of the height of the hip bone is less objectionable to the patient than that of the antero-posterior diameter, so that equations (*i*) and (*j*) possess a practical advantage over (*c*).

To facilitate reference all the equations for finding the obstetric conjugate are now collected together:

(*a*) $C = \cdot978D - 1\cdot553 \pm \cdot299.$

(*b*) $C = \cdot365A + 6\cdot081 \pm \cdot705.$

(*c*) $C = \cdot972D + \cdot023A - 1\cdot743 \pm \cdot298.$

(*d*) $C = \cdot972D + \cdot023A + \cdot001I - 1\cdot745 \pm \cdot299.$

(*e*) $C = \cdot110P + 10\cdot011 \pm \cdot738.$

(*f*) $C = \cdot989D - \cdot165P - 1\cdot229 \pm \cdot293.$

(*g*) $C = \cdot590R - 1\cdot037 \pm \cdot619.$

(*h*) $C = \cdot602L - 1\cdot310 \pm \cdot610.$

(*i*) $C = 1\cdot008D - \cdot049R - \cdot977 \pm \cdot298.$

(*j*) $C = 1\cdot005D - \cdot042L - 1\cdot075 \pm \cdot298.$

Equations (*f*), (*i*) and (*j*) appear to indicate that, for a constant diagonal conjugate, the obstetric conjugate is smaller the higher the pelvic wall.

These formulae have been tested for the original series and for five other pelves. The cards of 25 of the 216 pelves were picked out at random and the obstetric conjugate was calculated from the recorded measurements by means of the equations, with satisfactory results. The square root of the mean square deviation of the calculated from the recorded values of the obstetric conjugate was multiplied by ·67449 and compared with the probable errors of the equations. The values calculated from equations (*b*), (*g*), and (*h*) fell within this probable

error. The others exceeded it, but the greatest difference was only ·08, i.e., since we are working in centimetres, less than 1 millimetre.

The formulae were also tested for five pelves measured by us. Three of these were ancient Egyptian (I and II female, III male), and had to be fitted together to be measured, but the other two were modern English (IV female, V male), with the ligaments attached. The inter-tuberal diameter was not measured since it would have required the application of Emmons's method, and this we were not prepared to carry out as it had already been found that, in the determination of the obstetric conjugate and the transverse diameter, nothing was gained by the use of this diameter. Equation (*d*) was therefore not tested. The differences of the calculated from the recorded values for these pelves are given below for each equation.

TABLE IV.

Pelvis Number...	I	II	III	IV	V
Equation (*a*) ...	1·33	·39	– ·28	·15	– ·54
„ (*b*) ...	·83	– ·76	·52	2·65	·33
„ (*c*) ...	1·35	·40	– ·26	·19	– ·52
„ (*e*) ...	·44	– 1·09	·07	2·00	– ·15
„ (*f*) ...	1·26	·59	– ·01	·28	– ·32
„ (*g*) ...	1·49	– ·21	– ·51	2·24	– 1·35
„ (*h*) ...	1·43	– ·10	– ·53	2·42	– 1·43
„ (*i*) ...	1·27	·37	– ·23	·07	– ·45
„ (*j*) ...	1·29	·37	– ·23	·08	– ·46

Perhaps the most interesting of these numbers are those for the modern English female pelvis IV. Here the results for the equations formed from the diagonal conjugate and the height of the hip bone, (*i*) and (*j*), are excellent, the differences between recorded and calculated values being only ·07 and ·08 cm. The equations formed from the diagonal conjugate singly (*a*), and the diagonal conjugate and antero-posterior diameter (*c*), are the next best, the differences being ·15 and ·19 cm. respectively, then that from the diagonal conjugate and pubic height (*f*), with a difference ·28 cm. All the others have differences 2 cm. or more.

The figures bring out strongly what our previous discussion on the probable errors has shown, namely that, excluding, for reasons considered, the equations formed with the pubic height, the best equations are those containing the diagonal conjugate singly or combined with the height of the hip bone or the antero-posterior diameter.

The equations (*i*), (*j*), (*a*), (*c*) and (*f*) give differences less than a centimetre in four of the five pelves. No other equation does this except (*b*), but the difference for this in the remaining pelvis, which happens to be the modern English female pelvis IV, is 2·65 cm., so that the equations which give the best results on the whole are those which give the best results with pelvis IV. It

should be noted that all the equations give fairly good values for the male Egyptian pelvis III and most of them for the male English pelvis V.

The best of the formulae obtained from the measurements of the American Indian female pelvis for finding the obstetric conjugate are therefore applicable to the modern European female pelvis, and will give good results even for the male pelvis.

(B) *Transverse diameter (brim).* The inter-cristal (·61), inter-spinous (·52), and antero-posterior (·35), are the diameters most closely correlated with the transverse diameter. By taking them singly and combined, equations for finding this diameter were formed. Equations were also formed from the height of the hip bone taken singly or combined with the diameters. This bone measurement has proved serviceable in finding the obstetric conjugate and was used here, as its correlation with the transverse diameter also is fairly high (right ·48, left ·45). The following equations were obtained:

$$(k)\quad T = \cdot305K + 5\cdot096 \pm \cdot386.$$

$$(l)\quad T = \cdot228S + 7\cdot791 \pm \cdot415.$$

$$(m)\quad T = \cdot284A + 9\cdot673 \pm \cdot455.$$

$$(n)\quad T = \cdot261K + \cdot049S + 4\cdot864 \pm \cdot385.$$

$$(o)\quad T = \cdot221K + \cdot061S + \cdot160A + 4\cdot044 \pm \cdot374.$$

$$(p)\quad T = \cdot341R + 6\cdot398 \pm \cdot427.$$

$$(q)\quad T = \cdot320L + 6\cdot794 \pm \cdot433.$$

$$(r)\quad T = \cdot248K + \cdot144R + 3\cdot799 \pm \cdot378.$$

$$(s)\quad T = \cdot259K + \cdot115L + 4\cdot078 \pm \cdot381.$$

$$(t)\quad T = \cdot169S + \cdot221R + 4\cdot883 \pm \cdot392.$$

$$(u)\quad T = \cdot174S + \cdot192L + 5\cdot323 \pm \cdot398.$$

$$(v)\quad T = \cdot202K + \cdot050S + \cdot145R + 3\cdot835 \pm \cdot376.$$

$$(w)\quad T = \cdot217K + \cdot047S + \cdot113L + 4\cdot155 \pm \cdot380.$$

Of the equations formed by taking the measurements singly, namely (k) from inter-cristal, (l) inter-spinous, (m) antero-posterior diameter, (p) height of right hip bone, (q) height of left hip bone, (k) has the lowest probable error, ·386. This is slightly improved (·385) by combining the inter-spinous with the inter-cristal diameter as at (n). The inter-cristal diameter combined with the height of the hip bone gives a still lower probable error (right ·378, left ·381), as is seen in equations (r) and (s), while (t) and (u) show that the inter-spinous diameter and the height of the hip bone together give a higher probable error even than the inter-cristal diameter alone (k). By combining three measurements, however, there is noticeable improvement. Equation (v) formed from the inter-cristal and inter-spinous diameters and the height of the right hip bone has a probable error ·376 which is slightly lower than that of (r), from the inter-cristal diameter and height of right hip bone (·378); and (w), from the inter-cristal and inter-spinous

diameters and the height of the left hip bone, shows similar improvement as compared with (*s*), from the inter-cristal diameter and height of left hip bone (·381 and ·380 respectively); while (*o*), from the inter-cristal, inter-spinous, and antero-posterior diameters, has the lowest probable error, ·374.

The best formula then for finding the transverse diameter from a single diameter is (*k*), from the inter-cristal diameter. For two measurements, (*r*) and (*s*), from the inter-cristal diameter with the height of the right and left hip bones respectively, are slightly better than (*n*), from the inter-cristal and inter-spinous diameters, and any of these can conveniently be used as all the measurements are easily taken. In the combination of the inter-cristal and inter-spinous diameters with a third measurement, the antero-posterior diameter gives a better equation (*o*) than the height of the hip bone, (*v*) and (*w*), but the measurement of the latter is less objectionable to the patient.

These formulae were tested in the same way as those for the obstetric conjugate. The transverse diameter was calculated from the recorded measurements of 25 of the 216 pelves, picked out at random. The square root of the mean square deviation of the calculated from the recorded values of the transverse diameter was multiplied by ·67449 and compared with the probable errors of the equations. In every case they fell within this probable error.

The formulae were also tested for the five pelves measured by us (I and II Egyptian female, III Egyptian male, IV English female, V English male). The differences of the calculated from the recorded values for these are given below for each equation.

TABLE V.

Pelvis Number...	I	II	III	IV	V
Equation (*k*) ...	− ·58	−1·04	−1·18	− ·27	−1·57
" (*l*) ...	− ·99	− ·87	−1·58	+ ·17	−1·48
" (*m*) ...	−1·30	−1·13	− ·21	+ ·31	−1·08
" (*n*) ...	− ·33	− ·74	−1·06	+ ·08	−1·30
" (*o*) ...	− ·51	− ·93	−1·16	+ ·09	−1·40
" (*p*) ...	−1·00	− ·86	− ·86	− ·04	−2·12
" (*q*) ...	−1·09	− ·83	− ·83	+ ·05	−2·10
" (*r*) ...	− ·51	− ·90	−1·22	− ·20	−1·86
" (*s*) ...	− ·54	− ·91	−1·22	− ·18	−1·81
" (*t*) ...	− ·74	− ·77	−1·56	+ ·15	−1·95
" (*u*) ...	− ·81	− ·77	−1·45	+ ·20	−1·90
" (*v*) ...	− ·52	− ·87	−1·37	− ·12	−1·85
" (*w*) ...	− ·59	− ·91	−1·38	− ·13	−1·83

Two things are at once obvious, (1) that the best result for every equation is given by the modern English female pelvis IV, (2) that the results for the two male pelves III and V are not very good. In the case of the equations (*o*), (*v*), (*w*), (*r*), (*s*), (*n*) and (*k*), shown to have the least probable errors, the differences for the three female pelves, I, II, and IV, are all less than a centimetre except that of (*k*) for pelvis II, which is a little more. The best results for the English

female pelvis IV are given by (p) and (q), but their results are amongst the worst for pelvis I. The table shows that the equations selected from a consideration of probable errors may be expected to give good results, that these formulae may be applied to the modern European female pelvis, and that they cannot be satisfactorily applied to the male pelvis.

(4) *The effects of moisture on the pelvic measurements.*

The formulae which we have obtained will give the obstetric conjugate and the transverse diameter for the dry pelvis. These will not necessarily be the same in the living subject for there the bones contain moisture and are covered by the soft tissues. For practical purposes it is required to know what effect these differences will have on the measurements.

The effect of moisture was studied by soaking a dry pelvis in water and comparing its measurements in the dry state with those taken after it had been in water for some time. This was done with two ancient Egyptian female pelves lent by Dr Derry. The bones of the pelves were measured, then fitted together and kept in position by strong elastic bands, and the diameters measured with callipers. The pelvis was then taken to pieces and the bones immersed in water. After an interval they were taken out, held up to drip, and excess of water removed from them with a cloth. They were measured, fitted together again, and the diameters measured. This was repeated at several intervals. Finally they were taken out of the water and left in the air to dry for some days, and the bones and diameters were again measured. The results were unsatisfactory, as the following example, giving the measurements at different intervals, will show.

TABLE VI.

Pelvis I	Dry	In water 2 hours	In water 4 hours	In water 7 hours	In water 70 hours	In water 97 hours	After drying for 14 days
Inter-crests	22·30	22·52	22·45	22·39	22·82	23·02*	22·79
Inter-spines	19·81	19·95	19·99	19·91	20·23	20·52	20·23
Transverse	11·33	11·30	11·23	11·25	11·24	11·34	11·14
Diagonal conjugate ...	11·18	11·30	11·30	11·23	11·13	11·28	10·96
Obstetric conjugate ...	10·71	10·70	10·70	10·69	?	?	?
Antero-posterior ...	10·42	10·25	10·52	10·28	10·51	10·35	10·31
Pubic height	2·32	2·30	2·45	2·47	?	?	?
Right hip—height ...	17·40	17·50	17·47	17·50	17·50	17·53*	17·42
" " breadth ...	14·29	14·20	14·15	14·17	14·16	14·15	14·03
Left hip—height ...	17·60	17·75	17·60	17·60	17·60	17·63*	17·50
" " breadth ...	14·02	13·85	13·85	13·88	13·83	13·85	13·60
Sacrum—height ...	8·72	8·88	8·80	8·86	8·87	8·87	8·70
" breadth ...	10·30	10·16	10·17	10·13	10·13	10·14	10·07

Obviously the changes in the measurements are small and therefore not easy to detect where there are opportunities for error as in this case. The initial

* Both hip bones were cracked at the crest.

difficulty of putting the bones together on different occasions so as to get constant measurements is increased under the conditions of the experiment. After the bones have taken up water the shape of their articular surfaces is altered; hence the bones do not fit together in quite the same way as they did before, and the measurements are not fully comparable. A difference in the tilt of the sacrum, for example, may affect the two conjugates and the antero-posterior diameter.

But there is another difficulty. These old bones are very fragile and are apt to crack, as both hip bones did in this experiment. Flakes of bone come away from the surface. This happened at the upper part of the pubes in this pelvis, and is the reason why the record for the obstetric conjugate and pubic height is incomplete. Fortunately we were able to surmount these difficulties.

Prof. Thane lent us two modern English pelves, one female and one male (previously referred to as IV and V respectively), with the bones still attached to each other by the ligaments. Both were measured in the dry state. Pelvis IV was measured after immersion in water for 22 and 68 hours, and subsequent drying for 39 days. Pelvis V was measured after 98 hours in water, and subsequent drying for 35 days. The final measurements were very nearly the same as the initial, so that the mean of the two was taken as a good measure of the diameters in the dry pelvis. In the case of the first pelvis (IV) the numbers after 22 and 68 hours in water were also very close, and the mean of these was taken as the measurement in the moist state. The measurements in the dry state with their increase in the moist state are tabulated below.

The negative sign in the fourth column indicates that moisture caused some diminution of the measurement in question.

TABLE VII.

	Pelvis IV ♀		Pelvis V ♂		Mean of Differences
	Dry	Difference	Dry	Difference	
Inter-crests	26·26	·10	26·54	·18	·14
Inter-spines	21·40	·02	23·29	·04	·03
Transverse	15·85	·06	11·63	·08	·07
Diagonal conjugate ...	14·14	·15	12·70	·10	·13
Obstetric conjugate ...	12·42	·00	10·32	·02	·01
Antero-posterior ...	10·11	·13	10·71	·03	·08
Pubic height	3·72	·22	4·20	·21	·22
Right hip—height ...	19·03	·07	21·55	·05	·06
„ „ breadth ...	15·78	·06	15·87	– ·07	– ·01
Left hip—height ...	18·79	·11	21·70	·10	·11
„ „ breadth ...	15·77	·05	15·87	·03	·04

The differences are small, the greatest being a little over 2 mm. for the pubic height. Moreover they are very close for the two pelves. The means of these

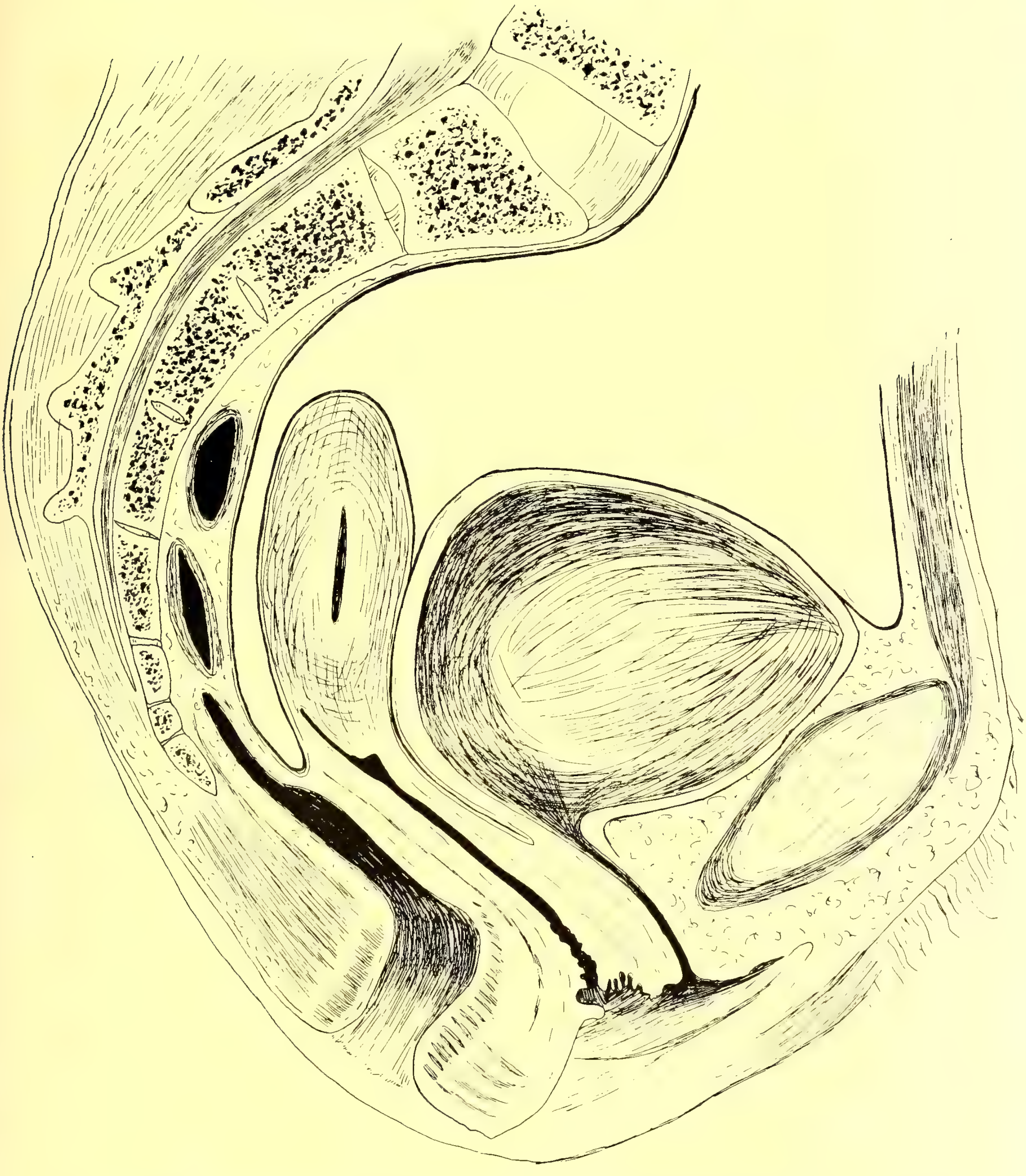


FIG. 1. Median section of the pelvis of a woman 35 years old. She was rather thin. From a drawing to scale by Professor Symington.

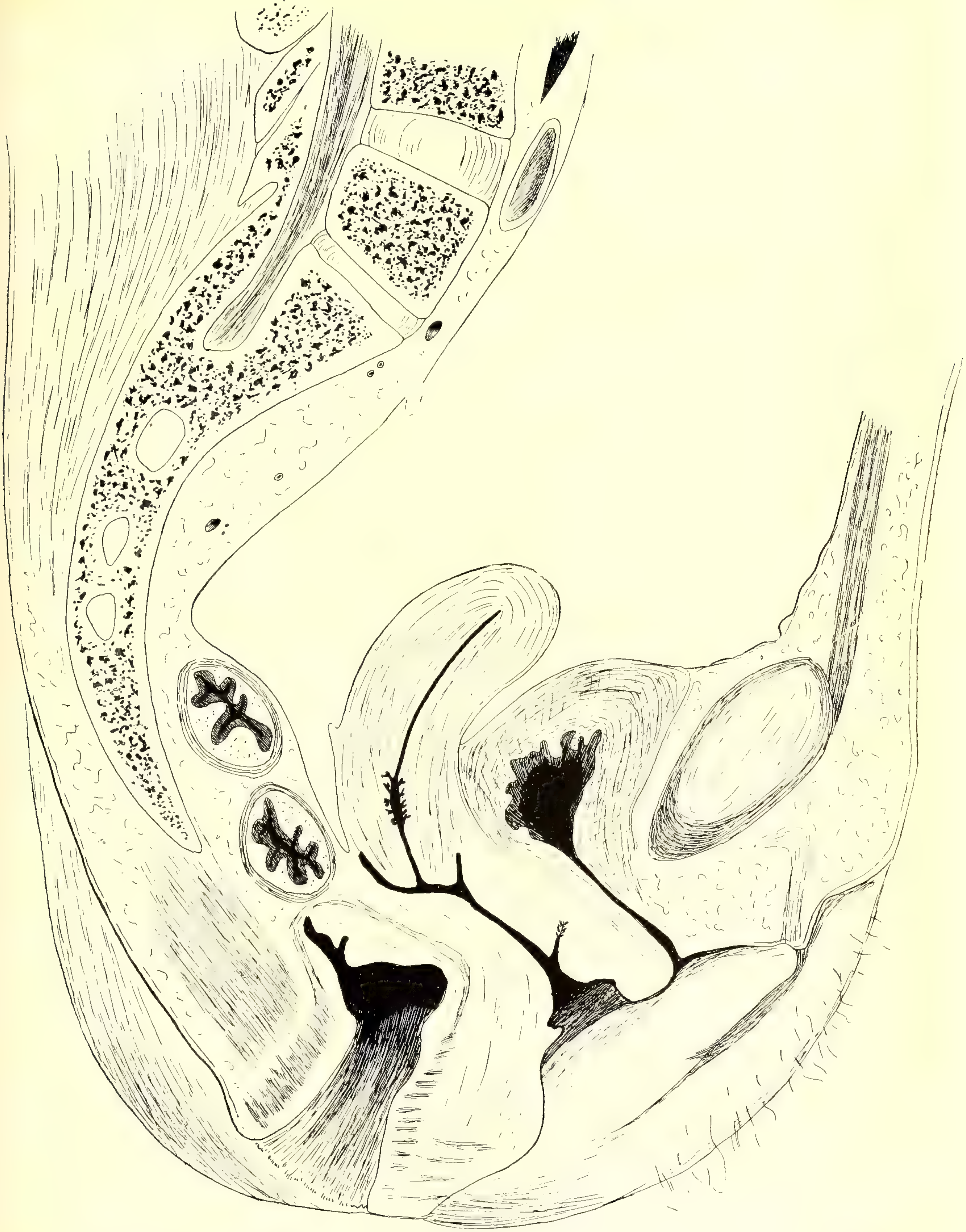


Fig. 2. Median section of the pelvis of a woman 30 years old. The subject was rather fat. From a drawing to scale by Professor Symington.

differences for the two pelves are given in the last column. In future calculations the numbers in this column are used to indicate the effect of moisture on the different measurements. It follows that, in the moist pelvis, all the diameters and the heights of the hip bones are slightly greater than in the dry pelvis.

(5) *The thickness of the soft parts.*

We have been unable to find any very satisfactory record of the thickness of the soft parts of the pelvis. The soft parts affect our calculations. An internal diameter such as the obstetric conjugate measured on the tissue-clothed pelvis must be less than it would be if measured on the moist bony pelvis, by the thickness of the tissues over the sacrum and pubes. An external measurement such as the height of the hip bone is greater on the living subject than on the moist pelvis by the thickness of the tissues above the iliac crest and over the ischium. The thickness of the tissues which increase the height of the hip bone as well as that of the tissues over the ilium in the inter-cristal diameter have been most kindly measured for us by Dr Derry on twenty adult female post-mortem subjects in the manner already described. The details of these measurements will be found in a table at the end of this paper. The averages only are given here.

TABLE VIII.

Inter-cristal diameter		28·63
Thickness of tissues in this diameter	right	·34
	left	·36
Ischio-iliac diameter	right	21·49
	left	21·62
Thickness of tissues *above* crest of ilium	right	·47
	left	·47
Thickness of tissues over ischium ...	right	·49
	left	·47

The inter-cristal diameter on the living subject is therefore greater than that of the moist bony pelvis by ·70 cm., that being the combined thickness of the tissues on the two sides. The ischio-iliac diameter, i.e. the height of the hip, is greater on the living subject by ·47 + ·49 or ·96 cm. on the right side, and ·47 + ·47 or ·94 cm. on the left side.

The tissues over the promontory of the sacrum and the posterior surface of the pubes were not measured. They are almost negligible. Two diagrams drawn to scale and most generously placed at our disposal by Prof. Symington show the differences for a thin and fat subject. In Fig. 1, Plate XXVIII, the median section of a rather thin woman 35 years old, the tissues over the promontory measure 2 mm. while those over the pubes cannot be measured on account of the distended bladder. In Fig. 2, Plate XXIX, showing the left half of the median section of the pelvis of a rather fat woman 30 years old, the tissues over the promontory measure 6 mm., over the pubes 4 mm., i.e. less than over the promontory.

These numbers, however, do not give a good idea of the thickness of the tissues in practice. The tissues would be compressed by the parts of the foetus, and, being composed of loose connective tissue, would take up much less room. If an attempt be made to measure them in a compressed condition the measurement is so small as to be almost negligible, and the chances of error are great. In what follows 1 mm. has been considered a just estimation of the thickness of the tissues over the surfaces of these bones under these conditions. The tissues will therefore diminish the diameters, as measured on the moist bony pelvis, by 2 mm. for the obstetric conjugate, diagonal conjugate and antero-posterior diameter, and will increase the pubic height by 1 mm., if measured internally.

With regard to the transverse diameter at the brim 2 mm. will also be a good allowance. Just at the brim, where this diameter is measured, there is nothing but loose tissue. The external iliac vessels and the psoas muscle are well above the brim, and it seems hardly likely that they can be pushed over it so as to diminish the transverse diameter.

(6) *The finding of the obstetric conjugate and the transverse diameter for the living subject.*

All the necessary information has now been collected for finding formulae to determine the obstetric conjugate and the transverse diameter in the living subject. These formulae can be obtained from the equations previously given by allowing for the effects of moisture and the soft parts. The same letters are used to indicate the lengths of the diameters, but are printed in clarendon to denote the length in the living subject. Consider the obstetric conjugate in the living subject. It is smaller than that of the moist bony pelvis by the thickness of the tissues over the bones, ·2 cm., hence, if this thickness be added to it, we get the obstetric conjugate of the moist bony pelvis. This again is larger than that of the dry pelvis by ·01 cm., the increase due to moisture. If ·01 cm. be subtracted, therefore, the obstetric conjugate is obtained for the dry pelvis. This is expressed by the equation

$$\mathbf{C} + \cdot 20 - \cdot 01 = C.$$

The same argument applies for the diagonal conjugate and antero-posterior and transverse diameters, the equations for which are

$$\mathbf{D} + \cdot 20 - \cdot 13 = D,$$

$$\mathbf{A} + \cdot 20 - \cdot 08 = A,$$

$$\mathbf{T} + \cdot 20 - \cdot 07 = T.$$

With the inter-spinous and inter-cristal diameters, the pubic height, and the height of the hip bone, the case is different. The thickness of the soft parts makes the last three greater in the living subject than on the moist bony pelvis and, to obtain them for the latter, this thickness must be subtracted. It does not, however, affect the inter-spinous diameter for which therefore no correction for soft parts is required. The measurements of the moist bony pelvis are, as with

the other diameters, larger than those of the dry pelvis, so that, to obtain these, the increase due to moisture must also be subtracted. The equations are

$$\mathbf{K} - {\cdot}70 - {\cdot}14 = K,$$
$$\mathbf{S} - \ \ 0 - {\cdot}03 = S,$$
$$\mathbf{P} - {\cdot}10 - {\cdot}22 = P,$$
$$\mathbf{R} - {\cdot}96 - {\cdot}06 = R,$$
$$\mathbf{L} - {\cdot}94 - {\cdot}11 = L.$$

If these values for C, D, A, T, K, S, P, R, and L be now substituted in the original formulae (a) to (w), the new formulae produced give the measurements on the living subject. In the following list these are also lettered (a') to (w') to indicate the original formula from which each was derived.

Obstetric conjugate.

(a') $\mathbf{C} = {\cdot}978\mathbf{D} - 1{\cdot}675 \pm {\cdot}299.$

(b') $\mathbf{C} = {\cdot}365\mathbf{A} + 5{\cdot}935 \pm {\cdot}705.$

(c') $\mathbf{C} = {\cdot}972\mathbf{D} + {\cdot}023\mathbf{A} - 1{\cdot}862 \pm {\cdot}298.$

(e') $\mathbf{C} = {\cdot}110\mathbf{P} + 9{\cdot}785 \pm {\cdot}738.$

(f') $\mathbf{C} = {\cdot}989\mathbf{D} - {\cdot}165\mathbf{P} - 1{\cdot}297 \pm {\cdot}293.$

(g') $\mathbf{C} = {\cdot}590\mathbf{R} - 1{\cdot}829 \pm {\cdot}619.$

(h') $\mathbf{C} = {\cdot}602\mathbf{L} - 2{\cdot}132 \pm {\cdot}610.$

(i') $\mathbf{C} = 1{\cdot}008\mathbf{D} - {\cdot}049\mathbf{R} - 1{\cdot}047 \pm {\cdot}298.$

(j') $\mathbf{C} = 1{\cdot}005\mathbf{D} - {\cdot}042\mathbf{L} - 1{\cdot}151 \pm {\cdot}298.$

Transverse diameter.

(k') $\mathbf{T} = {\cdot}305\mathbf{K} + 4{\cdot}710 \pm {\cdot}386.$

(l') $\mathbf{T} = {\cdot}228\mathbf{S} + 7{\cdot}654 \pm {\cdot}415.$

(m') $\mathbf{T} = {\cdot}284\mathbf{A} + 9{\cdot}577 \pm {\cdot}455.$

(n') $\mathbf{T} = {\cdot}261\mathbf{K} + {\cdot}049\mathbf{S} + 4{\cdot}513 \pm {\cdot}385.$

(o') $\mathbf{T} = {\cdot}221\mathbf{K} + {\cdot}061\mathbf{S} + {\cdot}160\mathbf{A} + 3{\cdot}747 \pm {\cdot}374.$

(p') $\mathbf{T} = {\cdot}341\mathbf{R} + 5{\cdot}920 \pm {\cdot}427.$

(q') $\mathbf{T} = {\cdot}320\mathbf{L} + 6{\cdot}328 \pm {\cdot}433.$

(r') $\mathbf{T} = {\cdot}248\mathbf{K} + {\cdot}144\mathbf{R} + 3{\cdot}313 \pm {\cdot}378.$

(s') $\mathbf{T} = {\cdot}259\mathbf{K} + {\cdot}115\mathbf{L} + 3{\cdot}610 \pm {\cdot}381.$

(t') $\mathbf{T} = {\cdot}169\mathbf{S} + {\cdot}221\mathbf{R} + 4{\cdot}522 \pm {\cdot}392.$

(u') $\mathbf{T} = {\cdot}174\mathbf{S} + {\cdot}192\mathbf{L} + 4{\cdot}986 \pm {\cdot}398.$

(v') $\mathbf{T} = {\cdot}202\mathbf{K} + {\cdot}050\mathbf{S} + {\cdot}145\mathbf{R} + 3{\cdot}387 \pm {\cdot}376.$

(w') $\mathbf{T} = {\cdot}217\mathbf{K} + {\cdot}047\mathbf{S} + {\cdot}113\mathbf{L} + 3{\cdot}723 \pm {\cdot}380.$

With this large number of formulae to choose from there should be no difficulty in getting a very close approximation to the lengths of these two non-measurable diameters. Since the desire to attain an end with the least possible exertion is a weakness of human nature, the majority of those who favour these formulae with a trial will, no doubt, wish to use a single measurement, and especially one from which both diameters can be found. Such measurements are the antero-posterior diameter in (b') and (m'), and the height of the hip bone, right in (g') and (p'), left in (h') and (q'). They must not, however, expect such good results as the man who measures the diagonal conjugate and the inter-cristal diameter to obtain the obstetric conjugate from (a) or, more approximately,

$$\mathbf{C} = \cdot 98\,\mathbf{D} - 1 \cdot 68,$$

and the transverse diameter from (k) or, more approximately,

$$\mathbf{T} = \cdot 31\,\mathbf{K} + 4 \cdot 71.$$

With a little extra labour the height of the hip bone can be added to the list. If $\mathbf{H}$ represent the height of either hip bone in the living subject equations (i') and (j') may be combined to give approximately

$$\mathbf{C} = \mathbf{D} - \cdot 05\,\mathbf{H} - 1,$$

or

$$\mathbf{C} = (\mathbf{D} - 1) - \tfrac{1}{20}\mathbf{H}.$$

Similarly (r') and (s') may be thrown into the combined form

$$\mathbf{T} = \tfrac{1}{4}(\mathbf{K} + 10) + \tfrac{1}{10}(\mathbf{H} + 10).$$

We thus arrive at the following simple rules:

(i) *To find the obstetric conjugate subtract one-twentieth of the height of either hip bone from the diagonal conjugate diminished by one centimetre.*

(ii) *To find the transverse diameter add one-quarter of the inter-cristal diameter increased by ten centimetres to one-tenth of the height of either hip bone increased by ten centimetres.*

(7) *On the application of the formulae to the pelves of dwarfs.*

It is important to know what reliance can be placed on these formulae in cases of extreme pelvic deformity. In an article on "Dwarfism" by Rischbieth and Barrington (*Treasury of Human Inheritance*, Vol. I, p. 355, 1912) several measurements of dwarfs of various types are recorded. In nine cases measurements of the pelvis are given which include the diagonal conjugate, but in only four of these is the obstetric conjugate also given. The obstetric conjugate has been calculated from the diagonal conjugate for these nine pelves by means of the equation

$$\mathbf{C} = \cdot 98\,\mathbf{D} + 1 \cdot 68$$

with the following results:

Number	Type of case	D	C (observed)	C (calculated)
1	Achondroplasia	7	—	5·18
2	„	9	—	7·14
3	Ateliosis	9·21	7·62	7·35
4	„	8·26	7·62	6·41
5	„	9·95	—	8·07
6	„	8·5	—	6·65
7	Rickety dwarf	7	6·25	5·18
8	„ „	8·75	—	6·90
9	Uncertain type	8·47	7·62	6·62

In No. 3 the calculated and measured values are very close, the former being slightly the less, but in 4, 7 and 9 the calculated value is about 1 cm. less than the observed value. In 4 the height of the hip bone was given, so that we were able to find the obstetric conjugate from the equation

$$\mathbf{C} = (\mathbf{D} - 1) - \tfrac{1}{20}\,\mathbf{H}.$$

The value obtained, 6·57, is a little nearer that of the observed value, 7·62, than is the value obtained from the diagonal conjugate alone, 6·41.

The last column shows that all the pelves are of such a degree of contraction as would call for special treatment during pregnancy or labour. The deformity would in most cases need no calculation to make it obvious. Even in extreme contraction, however, the formulae may be useful. In considering the advisability of Caesarian section in such cases, it is necessary to know whether the pelvis will admit the passage of even a dead child. This can be readily ascertained by applying the formulae. The tendency seems to be for them to give too low values in these extreme cases. With the present-day results of Caesarian section this is to err on the safe side for both the mother and the child.

SUMMARY.

1. Methods for obtaining the height of the hip bone on the living subject and for measuring the thickness of the tissues over different parts of this bone on the corpse, have been described.

2. The averages and variability of the pelvic measurements have been found and the various measurements correlated. The obstetric and diagonal conjugates show very high correlation. The inter-cristal is the diameter most closely correlated with the transverse diameter. Both the obstetric conjugate and the transverse diameter show good correlation with the height and breadth of the hip bone. Corresponding measurements on the two hip bones are more highly correlated than two measurements on the same bone.

3. Formulae have been obtained for finding the obstetric conjugate and the transverse diameter on the dry pelvis. The former may be found best from the diagonal conjugate, singly or combined with the antero-posterior diameter, the height of the hip bone, or the pubic height. The transverse diameter is found from the inter-cristal diameter, taken by itself or combined with one or two of the following: inter-spinous diameter, antero-posterior diameter, height of hip bone. The formulae were tested and found to give good results for the modern English female pelvis.

4. Pelves were measured dry and after immersion in water to study the effects of moisture. All the diameters and the heights of the hip bones are slightly greater in the moist than in the dry pelvis.

5. The way in which the soft tissues affect the measurements has been pointed out. The thickness of the tissues over the bones bounding the internal diameters is almost negligible.

6. Finally formulae have been arrived at for finding the obstetric conjugate and the transverse diameter in the living subject. The simplest of these, for the obstetric conjugate, are approximately

$$\mathbf{C} = \cdot 98\mathbf{D} - 1\cdot 68,$$

$$\mathbf{C} = (\mathbf{D} - 1) - \tfrac{1}{20}\mathbf{H},$$

and, for the transverse diameter,

$$\mathbf{T} = \cdot 31\mathbf{K} + 4\cdot 71,$$

$$\mathbf{T} = \tfrac{1}{4}(\mathbf{K} + 10) + \tfrac{1}{10}(\mathbf{H} + 10),$$

where **D** is the length of the diagonal conjugate, **K** that of the inter-cristal diameter, and **H** the height of either hip bone.

In conclusion I wish to thank Prof. Karl Pearson and all those who have so generously come to my help—Prof. Symington for placing his diagrams at my disposal, Prof. Thane and Dr Derry for the loan of pelves, Dr Derry and Dr Braxton Hicks for their measurements on post-mortem subjects, and Mr Soper for the useful little instrument with which most of these measurements were made. To Prof. Pearson is due whatever merit this work may possess. It has been carried out under his guidance without which it could never have been attempted.

APPENDIX I.

Measurements of five pelves used for testing formulae. I and II ancient Egyptian female. III ancient Egyptian male. IV modern English female. V modern English male. In IV and V the bones were connected together by the ligaments.

TABLE A.

Measurement	I	II	III	IV	V
Inter-crests	22·30	24·92	28·28	26·26	26·54
Inter-spines	19·81	21·25	27·75	21·40	23·29
Transverse	11·33	11·67	12·55	12·85	11·63
Diagonal conjugate ...	11·18	10·75	12·68	14·14	12·70
Obstetric conjugate ...	10·71	9·35	10·57	12·42	10·32
Antero-posterior ...	10·42	11·03	10·88	10·11	10·71
Pubic height	2·32	3·88	4·47	3·72	4·20
Right hip—height ...	17·40	17·98	20·55	19·03	21·55
,, ,, breadth ...	14·29	14·32	15·41	15·78	15·87
Left hip—height ...	17·60	17·87	20·62	18·79	21·70
,, ,, breadth ...	14·02	14·49	15·80	15·77	15·87
Sacrum—height ...	8·72	9·35	10·93	—	—
,, breadth ...	10·30	10·85	11·83	—	—

Table of measurements taken by Dr Derry from 20 adult female post-mortem subjects to show the thickness of the soft parts.

TABLE B.

Age	Inter-cristal diameter	Thickness of tissues in this diameter		Ischio-iliac diameter		Thickness of tissues *above* crest of ilium		Thickness of tissues over ischium		Remarks
		R	*L*	*R*	*L*	*R*	*L*	*R*	*L*	
61	31·15	·65	·79	19·95	20·00	·45	·55	·50	·55	not very fat but well nourished
35	26·35	·48	·49	19·75	19·60	·55	·60	·50	·40	,, ,, ,,
18	28·50	·45	·40	20·00	19·90	·39	·40	·60	·46	,, ,, ,,
56	29·00	·55	·50	23·00	23·00	·80	·80	·50	·45	,, ,, ,,
57	28·50	·55	·60	23·20	23·55	1·00	·85	·55	·45	stout
54	25·00	·40	·40	20·30	20·35	·55	·45	·50	·50	slight layer of fat only
55	28·95	·35	·40	21·15	21·65	·50	·50	·50	·45	,, ,, ,,
62	32·70	·30	·30	22·95	23·00	·40	·45	·50	·70	thin
31	26·95	·20	·10	20·50	20·40	·30	·20	·40	·45	*very* emaciated
65	28·50	·20	·35	21·00	21·20	·30	·45	·25	·30	thin
28	28·75	·20	·20	21·20	21·40	·35	·35	·40	·30	emaciated
59	30·30	·35	·35	23·00	23·00	·35	·35	·40	·50	thin
20	28·50	·30	·30	21·70	21·90	·60	·50	·60	·55	thin layer of fat
55	27·30	·40	·50	22·00	22·50	·50	·55	·60	·65	fat
39	29·10	·30	·25	21·05	21·35	·45	·40	·50	·45	thin
28	30·30	·35	·35	22·35	22·40	·35	·40	·40	·30	fairly fat
38	30·25	·20	·30	22·45	22·85*	·40	·45	·55	—*	thin
45	27·10	·10	·20	22·30	22·25	·45	·40	·60	·60	,,
22	26·10	·20	·15	21·25	21·35	·25	·25	·60	·55	,,
43	29·30	·20	·20	20·75	20·75	·40	·40	·40	·40	,,

* Some swelling over ischium.

APPENDIX II.

CORRELATION TABLES OF THE DIAMETERS WITH ONE ANOTHER.

TABLE I. Inter-cristal.

Obstetric conjugate.	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	27·25—	27·75—	28·25—	28·75—	Totals
7·25—	—	—	—	—	—	—	—	1	0·5	—	—	—	—	—	—	—	—	1·5
7·75—	—	—	—	—	—	—	—	—	0·5	—	1	2	—	—	—	—	—	3·5
8·25—	1	—	1	—	1	—	—	2	—	—	1	0·5	0·5	—	—	—	—	7
8·75—	—	—	—	—	1	1	3·5	1·5	6	3·5	6·5	—	4	0·5	1	1	—	29·5
9·25—	—	—	1	—	1	—	2	4	2·5	3	3	5	1·5	1·5	—	—	—	24·5
9·75—	—	—	—	—	1·5	—	3	1·5	5	6·5	6	2·5	5·75	1·25	4	3	—	40
10·25—	—	—	1	—	0·5	0·25	2·25	0·5	5·5	4	5	3	4·25	3·75	2	—	—	32
10·75—	—	—	—	—	3	0·25	5·25	3	3·5	3·5	6	1·5	5·25	1·75	4	1	—	38
11·25—	—	—	—	—	—	0·25	0·25	—	3	1	3	1	6·25	2·25	—	1	—	18
11·75—	—	—	—	—	—	1·25	0·25	3	2	2	2·5	2	—	0·5	0·5	—	1	15
12·25—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	—	—	2
12·75—	—	—	—	—	—	—	—	—	1	1	0·5	0·5	1	—	—	—	—	4
13·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
13·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	1
Totals	1	0	3	0	8	3	16·5	16·5	30·5	24·5	34·5	18	28·5	12·5	12·5	6	1	216

TABLE II. Transverse.

Obstetric conjugate.	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
7·25—	—	—	—	1	—	0·5	—	—	—
7·75—	—	—	—	—	1	2·5	—	—	—
8·25—	1	—	—	1	1	3	1	—	—
8·75—	—	—	1	6	4	11	3	4·5	—
9·25—	—	1	1	2	5·5	8	3·5	3	0·5
9·75—	—	—	1	3·5	4·5	13	9	7·75	1·25
10·25—	—	—	0·25	2·75	5·5	12	5	4·25	2·25
10·75—	1	—	3·25	8·75	4·5	9	7·5	2	2
11·25—	—	—	—	1	3	8·5	3·5	1	1
11·75—	—	—	—	2	0·5	6·5	5	1	—
12·25—	—	—	—	—	—	1	—	1	—
12·75—	—	—	—	—	—	4	—	—	—
13·25—	—	—	—	—	—	—	—	—	—
13·75—	—	—	—	1	—	—	—	—	—
Totals	2	1	6·5	29	29·5	79	37·5	24·5	7

TABLE III. Inter-spinous.

Obstetric conjugate.	16·75—	17·25—	17·75—	18·25—	18·75—	19·25—	19·75—	20·25—	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	Totals
7·25—	—	—	—	—	—	—	—	—	—	—	1	—	—	0·5	—	—	—	—	—	—	—	1·5
7·75—	—	—	—	—	—	—	—	—	—	—	0·5	0·5	1	0·5	1	—	—	—	—	—	—	3·5
8·25—	1	—	—	—	1	—	—	—	1·5	0·5	1	—	—	1·5	0·5	—	—	—	—	—	—	7
8·75—	—	—	—	—	1	2	1	—	2	2·5	2	6	4	3	3	1	2	—	—	—	—	29·5
9·25—	0·5	0·5	—	—	—	—	1	2	1	0·5	6	3	2	5·5	0·5	1	1	—	—	—	—	24·5
9·75—	—	—	—	—	—	—	1	2·25	4·25	2	5·25	1·25	4	5	6·25	3·75	1·5	2·5	1	—	—	40
10·25—	—	—	—	—	—	—	1·5	1·25	1·25	2	8·25	0·75	3	2	4·25	3·75	2·5	1·5	—	—	—	32
10·75—	—	—	—	—	1	2	3·5	2	1	4	7	1·5	3·5	3	7	—	0·25	1·25	—	—	1	38
11·25—	—	—	—	—	—	—	—	0·25	1·25	2	1·5	1·5	2·5	1·25	2·75	1·5	0·25	1·25	0·5	1·5	—	18
11·75—	—	—	—	—	—	—	—	2·25	1·25	1·5	3·5	2	2	0·25	1·25	—	1	—	—	—	—	15
12·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	—	—	—	—	—	—	2
12·75—	—	—	—	—	—	—	—	—	—	—	—	—	2	1	—	1	—	—	—	—	—	4
13·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
13·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	1·5	0·5	0	0	3	4	8	10	13·5	15	36	16·5	24	23·5	28·5	12	9·5	6·5	1·5	1·5	1	216

TABLE IV. Diagonal conjugate. TABLE V. Antero-posterior.

Obstetric conjugate.

	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	Totals
7·25—	1	0·5	—	—	—	—	—	—	—	—	—	—	1·5
7·75—	1	1·5	1	—	—	—	—	—	—	—	—	—	3·5
8·25—	—	3	2	2	—	—	—	—	—	—	—	—	7
8·75—	—	1	2	15	9	1·5	1	—	—	—	—	—	29·5
9·25—	—	—	1	5	14	4·5	—	—	—	—	—	—	24·5
9·75—	—	—	—	4	15·5	14	5	0·5	1	—	—	—	40
10·25—	—	—	—	—	1·75	13·5	14·25	2·5	—	—	—	—	32
10·75—	—	—	—	—	0·25	7·5	14·75	10·5	5	—	—	—	38
11·25—	—	—	—	—	—	—	0·5	10·75	3·75	2	1	—	18
11·75—	—	—	—	—	—	—	—	3·25	7·75	2	2	—	15
12·25—	—	—	—	—	—	—	—	—	2	—	—	—	2
12·75—	—	—	—	—	—	—	—	—	—	1	2	1	4
13·25—	—	—	—	—	—	—	—	—	—	—	—	—	0
13·75—	—	—	—	—	—	—	—	—	—	—	—	1	1
Totals	2	6	6	26	40·5	41	35·5	27·5	19·5	5	5	2	216

	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
7·25—	—	—	—	—	0·5	1	—	—	—	—	—	—
7·75—	—	—	—	—	1·5	1	1	—	—	—	—	—
8·25—	—	1	2	2	2	—	—	—	—	—	—	—
8·75—	—	—	1	7	5	9·5	4	3	—	—	—	—
9·25—	—	—	—	4	5	4	6·5	2	—	1·5	0·5	1
9·75—	1	—	1·5	3	12	4·5	8	8	2	—	—	—
10·25—	—	—	1·5	1·5	7·5	4·5	10	4	2	1	—	—
10·75—	—	1	1	6	10·5	5·5	3	6	4	—	1	—
11·25—	—	—	—	1	2·5	5	3·5	2	3	1	—	—
11·75—	—	—	—	—	1·5	5	2·5	1	4	1	—	—
12·25—	—	—	—	—	—	—	—	1	—	1	—	—
12·75—	—	—	—	—	1	—	1	1·5	0·5	—	—	—
13·25—	—	—	—	—	—	—	—	—	—	—	—	—
13·75—	—	—	—	—	—	—	—	—	1	—	—	—
Totals	1	2	7	24·5	49	40	39·5	28·5	16·5	5·5	1·5	1

TABLE VI. Inter-tuberal. TABLE VII. Posterior sagittal.

Obstetric conjugate.

	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
7·25—	—	—	1	—	0·5	—	—	—	—	—	—	1·5
7·75—	—	2	1	—	0·5	—	—	—	—	—	—	3·5
8·25—	—	2	1	2	1	—	—	1	—	—	—	7
8·75—	4	5	1·5	6	6	5	2	—	—	—	—	29·5
9·25—	—	2	6·5	5	5	5	—	—	—	0·5	0·5	24·5
9·75—	—	5	6·5	7·5	9	7	3	1	—	1	—	40
10·25—	—	1	6	3·5	7·5	5	7	2	—	—	—	32
10·75—	2	2	9	4	12	4	4	1	—	—	—	38
11·25—	2·5	—	1	2	5·5	5	—	1·5	0·5	—	—	18
11·75—	0·5	1·5	4	2	1	5	1	—	—	—	—	15
12·25—	—	—	—	—	—	—	2	—	—	—	—	2
12·75—	—	—	—	1	2	—	—	—	1	—	—	4
13·25—	—	—	—	—	—	—	—	—	—	—	—	0
13·75—	—	—	—	—	—	1	—	—	—	—	—	1
Totals	9	20·5	37·5	33	50	37	19	6·5	1·5	1·5	0·5	216

	5·75—	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—
7·25—	—	0·5	—	—	1	—	—	—	—	—	—	—
7·75—	—	1·5	—	—	2	—	—	—	—	—	—	—
8·25—	1	1	3	1	—	1	—	—	—	—	—	—
8·75—	2	8	8	6·5	2	2	—	1	—	—	—	—
9·25—	1	5	6	3·5	5	2	1	—	—	1	—	—
9·75—	—	1·5	12·5	14	5	4	2	1	—	—	—	—
10·25—	1	4·5	4·5	7·5	7·5	4	2	—	—	—	—	1
10·75—	1	5	4·5	10·5	11·5	1·5	2	2	—	—	—	—
11·25—	—	—	3	7	2·5	2·5	3	—	—	—	—	—
11·75—	—	1	4·5	2	5	1·5	1	—	—	—	—	—
12·25—	—	1	—	—	—	1	—	—	—	—	—	—
12·75—	—	—	1	1	1	1	—	—	—	—	—	—
13·25—	—	—	—	—	—	—	—	—	—	—	—	—
13·75—	—	—	—	—	1	—	—	—	—	—	—	—
Totals	6	29	47	53	43·5	20·5	11	4	0	1	0	1

TABLE VIII. Obstetric conjugate.

Pubic height.	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	Totals
0·75—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	2
1·25—	—	—	—	—	1	—	1	1	—	—	—	—	—	—	3
1·75—	—	—	1	3	2	4	6	4	1	2	—	—	—	—	23
2·25—	0·5	1·5	3	6	7	11·5	9·5	12	5	2	—	3·5	—	—	61·5
2·75—	1	1	2	18·5	9·5	18	7	11	8	3·5	2	0·5	—	1	83
3·25—	—	1	—	1	3	4·5	7	9·5	4	4·5	—	—	—	—	34·5
3·75—	—	—	—	1	1	1	—	—	—	1	—	—	—	—	4
4·25—	—	—	—	—	—	—	0·5	0·5	—	—	—	—	—	—	1
Totals	1·5	3·5	6	29·5	23·5	40	32	38	18	13	2	4	0	1	212

TABLE IX. Transverse.

Pubic height.	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
0·75—	—	—	—	2	—	—	—	—	—
1·25—	1	—	—	—	—	0·5	1·5	—	—
1·75—	—	—	2	3	1	10	4·5	2·5	—
2·25—	1	1	—	6·5	10·5	28·5	5	8	1
2·75—	—	—	2	12	12·75	27·75	17·5	9·5	1·5
3·25—	—	—	2	4	4·75	9·75	6	4	4
3·75—	—	—	—	1	0·5	1·5	1	—	—
4·25—	—	—	0·5	0·5	—	—	—	—	—
Totals	2	1	6·5	29	29·5	78	35·5	24	6·5

TABLE X. Inter-cristal.

Pubic height.	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	27·25—	27·75—	28·25—	28·75—	Totals
0·75—	—	—	1	—	—	—	—	1	—	—	—	—	—	—	—	—	—	2
1·25—	—	—	—	—	1	—	—	—	—	—	—	—	—	2	—	—	—	3
1·75—	—	—	—	—	—	1	4	2	7·5	3	2·5	1	2	—	—	—	—	23
2·25—	1	—	—	—	2	1	8	2·5	13	7	12·5	4	6·5	2	2	—	—	61·5
2·75—	—	—	2	—	4	—	3·5	5·75	7·25	8·5	15	6·5	16	5·5	5	4	—	83
3·25—	—	—	—	—	1	0·5	0·5	3·25	2·75	5	4	4·5	3·5	2·5	4	2	1	34·5
3·75—	—	—	—	—	—	—	—	1	—	—	—	1	—	0·5	1·5	—	—	4
4·25—	—	—	—	—	—	0·5	0·5	—	—	—	—	—	—	—	—	—	—	1
Totals	1	0	3	0	8	3	16·5	15·5	30·5	23·5	34	17	28	12·5	12·5	6	1	212

TABLE XI. Inter-spinous.

Pubic height.	16·75—	17·25—	17·75—	18·25—	18·75—	19·25—	19·75—	20·25—	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	Totals
0·75—	—	—	—	—	—	—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	2
1·25—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	3
1·75—	—	—	—	—	1	1	—	3	4	2	4	1	5	1	—	1	—	—	—	—	—	23
2·25—	1	—	—	—	1	1	2	4	3·5	2·5	12·5	7	8·5	8	6·5	3	1	—	—	—	—	61·5
2·75—	0·5	0·5	—	—	1	—	3	0·5	2·5	7	12·5	5	9·5	11	15·5	5·5	4	4	0·5	0·5	—	83
3·25—	—	—	—	—	—	1	1	1·5	1·5	3·5	5	2·5	1	3	5	1·5	3·5	1·5	1	1	1	34·5
3·75—	—	—	—	—	—	—	—	—	1	—	—	1	—	—	1	1	—	—	—	—	—	4
4·25—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
Totals	1·5	0·5	0	0	3	4	8	9	12·5	15	35	16·5	24	23	28	12	9·5	6·5	1·5	1·5	1	212

TABLE XII. Diagonal conjugate.

Pubic height.	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	Totals
0·75—	—	—	—	—	1	1	—	—	—	—	—	—	2
1·25—	—	—	—	—	1	—	1	1	—	—	—	—	3
1·75—	—	1	1	1	5	8	3	1	3	—	—	—	23
2·25—	—	2	5	10	9	12	9	5	6	1	1·5	1	61·5
2·75—	2	2	—	13	17	14·5	13	12·5	4·5	2	1·5	1	83
3·25—	—	1	—	1	4	4	9·5	7	6	1	1	—	34·5
3·75—	—	—	—	—	2	1	—	—	—	—	1	—	4
4·25—	—	—	—	—	0·5	0·5	—	—	—	—	—	—	1
Totals	2	6	6	25	39·5	41	35·5	26·5	19·5	4	5	2	212

TABLE XIII. Antero-posterior.

Pubic height.	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—
0·75—	—	—	—	—	1	1	—	—	—	—	—
1·25—	—	—	—	—	—	1	1	—	—	0·5	0·5
1·75—	—	—	2	4	2·5	4·5	7	2	1	—	—
2·25—	—	—	2	11·5	13	9	11	9·75	3·25	1	1
2·75—	1	2	1·5	8·5	20	16·5	11·5	12·75	7·25	2	—
3·25—	—	—	1·5	0·5	8·5	7	8	3	5	1	—
3·75—	—	—	—	—	1	1	—	1	—	1	—
4·25—	—	—	—	—	1	—	—	—	—	—	—
Totals	1	2	7	24·5	47	40	38·5	28·5	16·5	5·5	1·5

TABLE XIV. Inter-tuberal.

Pubic height.	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	Totals
0·75—	—	—	—	1	—	1	—	—	—	—	2
1·25—	—	1	—	—	1	—	1	—	—	—	3
1·75—	—	2	3	5	6	5	2	—	—	—	23
2·25—	3	4	11	6	19·5	7	8	2	1	—	61·5
2·75—	4	11	14	14	17	15	6	1·5	0·5	—	83
3·25—	1	2	8	6	5·5	7	2	2	—	1	34·5
3·75—	1	0·5	0·5	1	—	1	—	—	—	—	4
4·25—	—	—	—	—	1	—	—	—	—	—	1
Totals	9	20·5	36·5	33	50	36	19	5·5	1·5	1	212

TABLE XV. Posterior sagittal.

Pubic height.	5·75—	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—
0·75—	—	—	1	1	—	—	—	—	—	—	—	—
1·25—	—	—	—	2	—	—	—	—	—	1	—	—
1·75—	1	1	7	8	4	1	1	—	—	—	—	—
2·25—	2	11·5	8·5	13	10·5	10	2	4	—	—	—	—
2·75—	3	9·5	22	23·5	16	6	2	—	—	—	—	—
3·25—	—	5	6·5	5	10	3	5	—	—	—	—	1
3·75—	—	1	1	0·5	1	0·5	—	—	—	—	—	—
4·25—	—	—	—	—	1	—	—	—	—	—	—	—
Totals	6	28	46	53	42·5	20·5	10	4	0	1	0	1

TABLE XVI. Inter-cristal.

Transverse.	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	27·25—	27·75—	28·25—	28·75—	Totals
10·25—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	2
10·75—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	1
11·25—	—	—	1	—	—	0·25	2·25	1	1	1	—	—	—	—	—	—	—	6·5
11·75—	—	—	2	—	3	0·25	6·25	7	5·25	0·25	1	1	1	1	1	—	—	29
12·25—	—	—	—	—	2	—	3·5	3·75	1	4·25	5·5	3·5	4·5	—	1·5	—	—	29·5
12·75—	—	—	—	—	1	2·5	3·5	2·25	18·75	11	16·25	7·75	12	1·5	2·5	—	—	79
13·25—	—	—	—	—	—	—	1	2	4	4·5	7·25	3·25	5·5	5·5	1·5	3	—	37·5
13·75—	—	—	—	—	—	—	—	0·5	0·5	3·5	4·25	2·25	5·25	1·25	3	3	1	24·5
14·25—	—	—	—	—	—	—	—	—	—	—	0·25	0·25	0·25	3·25	3	—	—	7
Totals	1	0	3	0	8	3	16·5	16·5	30·5	24·5	34·5	18	28·5	12·5	12·5	6	1	216

TABLE XVII. Inter-spinous.

Transverse.	16·75—	17·25—	17·75—	18·25—	18·75—	19·25—	19·75—	20·25—	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	Totals
10·25—	1	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
10·75—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	1
11·25—	0·5	0·5	—	—	1	1	0·5	1	—	1	1	—	—	—	—	—	—	—	—	—	—	6·5
11·75—	—	—	—	—	—	—	4·5	0·5	5·5	3·5	6·75	2·25	2	1	2	—	1	—	—	—	—	29
12·25—	—	—	—	—	—	1	1	1	1·5	2·5	7·25	2·25	0·5	6·75	3·5	1·25	—	1	—	—	—	29·5
12·75—	—	—	—	—	2	1	1	4	2·5	6	13	8·75	13·75	9·75	8·5	7·25	1	0·5	—	—	—	79
13·25—	—	—	—	—	—	—	1	2	2	1	6	3·25	5·25	2·5	6	1·5	2	4	—	—	1	37·5
13·75—	—	—	—	—	—	—	—	0·5	1·5	1	2	—	2·5	2·5	7·5	2	3·5	0·5	0·5	0·5	—	24·5
14·25—	—	—	—	—	—	—	—	—	0·5	—	—	—	—	1	1	—	2	0·5	1	1	—	7
Totals	1·5	0·5	0	0	3	4	8	10	13·5	15	36	16·5	24	23·5	28·5	12	9·5	6·5	1·5	1·5	1	216

TABLE XVIII. Diagonal conjugate.

Transverse.	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	Totals
10·25—	—	—	1	—	—	—	—	1	—	—	—	—	2
10·75—	—	—	—	—	—	1	—	—	—	—	—	—	1
11·25—	—	—	—	2	0·25	2·25	1	1	—	—	—	—	6·5
11·75—	1	1	—	3	7·25	3·25	5·5	4	3	—	—	1	29
12·25—	—	1	2	3	6·5	5·5	5	6	0·5	—	—	—	29·5
12·75—	1	4	2	9	15	13·5	9	8·5	12	2	2	1	79
13·25—	—	—	—	6	6	6·5	8	5	1	3	2	—	37·5
13·75—	—	—	1	3	5	6·5	5	1	2	—	1	—	24·5
14·25—	—	—	—	—	0·5	2·5	2	1	1	—	—	—	7
Totals	2	6	6	26	40·5	41	35·5	27·5	19·5	5	5	2	216

TABLE XIX. Antero-posterior.

Transverse.	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
10·25—	—	—	0·5	0·5	—	1	—	—	—	—	—	—
10·75—	—	—	—	1	—	—	—	—	—	—	—	—
11·25—	—	—	2	1	2·5	1	—	—	—	—	—	—
11·75—	—	1	0·5	3	8·5	9	5	—	2	—	—	—
12·25—	—	—	0·5	7	3·75	4·75	6	6·5	—	1	—	—
12·75—	0·5	1	3·5	8·5	20·25	15·75	10·5	13	4·5	1·25	0·25	—
13·25—	0·5	—	—	2	8·5	5·5	10	5	4·5	1·25	0·25	—
13·75—	—	—	—	1·5	5	3	8	1	3·5	1	1	0·5
14·25—	—	—	—	—	0·5	—	—	3	2	1	—	0·5
Totals	1	2	7	24·5	49	40	39·5	28·5	16·5	5·5	1·5	1

TABLE XX. Inter-tuberal.

Transverse.	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
10·25—	—	1	—	—	—	—	1	—	—	—	—	2
10·75—	—	—	—	—	—	1	—	—	—	—	—	1
11·25—	—	2	1	—	2·5	—	1	—	—	—	—	6·5
11·75—	3	6	5·5	5	5·5	3	1	—	—	—	—	29
12·25—	3	3	6·5	7·5	4	1·5	4	—	—	—	—	29·5
12·75—	2	5	15	11·5	24	13·5	6	1	1	—	—	79
13·25—	1	1·5	8	4	7	11	2	2·5	0·5	—	—	37·5
13·75—	—	2	1·5	5	6	4·5	3	2	—	0·25	0·25	24·5
14·25—	—	—	—	—	1	2·5	1	1	—	1·25	0·25	7
Totals	9	20·5	37·5	33	50	37	19	6·5	1·5	1·5	0·5	216

TABLE XXI. Posterior sagittal.

Transverse.	5·75—	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—
10·25—	—	—	1	1	—	—	—	—	—	—	—	—
10·75—	1	—	—	—	—	—	—	—	—	—	—	—
11·25—	—	1	4	—	1·5	—	—	—	—	—	—	—
11·75—	1	4	9	6·5	6·5	1	—	1	—	—	—	—
12·25—	1	5	4·5	6·75	5·75	4·5	—	1	—	—	—	1
12·75—	2	12	19	16·25	14·75	8·5	5	1	—	0·5	—	—
13·25—	1	3	4·5	14	9	2·5	3	—	—	0·5	—	—
13·75—	—	3·5	5	7·5	3·5	3	1	1	—	—	—	—
14·25—	—	0·5	—	1	2·5	1	2	—	—	—	—	—
Totals	6	29	47	53	43·5	20·5	11	4	0	1	0	1

TABLE XXII. Diagonal conjugate.

Inter-spinous.	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	Totals
16·75—	—	—	1	0·5	—	—	—	—	—	—	—	—	1·5
17·25—	—	—	—	0·5	—	—	—	—	—	—	—	—	0·5
17·75—	—	—	—	—	—	—	—	—	—	—	—	—	0
18·25—	—	—	—	—	—	—	—	—	—	—	—	—	0
18·75—	—	—	1	1	—	1	—	—	—	—	—	—	3
19·25—	—	—	—	1	1	—	—	2	—	—	—	—	4
19·75—	—	—	—	—	3·5	2·5	2	—	—	—	—	—	8
20·25—	—	—	—	—	2	2	3	0·5	1·5	1	—	—	10
20·75—	—	1·5	—	1	5	1	1	2·5	0·5	—	1	—	13·5
21·25—	—	0·5	—	1	1	4	3	3	1·5	—	1	—	15
21·75—	1	1	1·5	3	6	8·5	5·5	5	4·5	—	—	—	36
22·25—	—	—	0·5	5	5	1·5	1	1·5	2	—	—	—	16·5
22·75—	1	1	1	2	3	5·5	3·5	1·5	2·5	2	1	—	24
23·25—	—	1	1	5·5	5·5	4	3·5	0·25	1·75	—	—	1	23·5
23·75—	—	1	—	1·5	4·5	4·5	7	5·75	4·25	—	—	—	28·5
24·25—	—	—	—	1	3	3	1·5	1·5	—	1	1	—	12
24·75—	—	—	—	2	1	1·5	2	—	1	—	1	1	9·5
25·25—	—	—	—	1	—	1	2·5	1	—	1	—	—	6·5
25·75—	—	—	—	—	—	1	—	0·5	—	—	—	—	1·5
26·25—	—	—	—	—	—	—	—	1·5	—	—	—	—	1·5
26·75—	—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	2	6	6	26	40·5	41	35·5	27·5	19·5	5	5	2	216

TABLE XXIII. Antero-posterior.

Inter-spinous.	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
16·75—	—	—	0·5	0·5	0·5	—	—	—	—	—	—	—
17·25—	—	—	—	—	0·5	—	—	—	—	—	—	—
17·75—	—	—	—	—	—	—	—	—	—	—	—	—
18·25—	—	—	—	—	—	—	—	—	—	—	—	—
18·75—	—	—	2	—	1	—	—	—	—	—	—	—
19·25—	—	—	—	—	1	2	1	—	—	—	—	—
19·75—	—	—	—	3	3	2	—	—	—	—	—	—
20·25—	—	—	0·25	2·25	2	1·5	2	1	1	—	—	—
20·75—	—	—	0·5	2·5	2	1·5	3·5	1·5	—	1	—	1
21·25—	—	1	0·25	1·25	1	5·5	2	3	1	—	—	—
21·75—	—	—	1	5·5	10	8·5	6	2	2	1	—	—
22·25—	—	—	1	2·5	2	4·5	2	1·5	2	1	—	—
22·75—	—	—	—	2	5·5	4·5	7·25	3·75	1	—	—	—
23·25—	—	1	—	2	5·5	5·5	4·25	4·25	1	—	—	—
23·75—	—	—	0·25	2·5	8·5	2·75	4	6·5	2	1	1	—
24·25—	—	—	1·25	0·5	3·5	1·25	2	2	1·5	—	—	—
24·75—	—	—	—	—	1·5	0·5	2·5	2	2	0·5	0·5	—
25·25—	1	—	—	—	1·5	—	2	1	1	—	—	—
25·75—	—	—	—	—	—	—	—	—	1·5	—	—	—
26·25—	—	—	—	—	—	—	—	—	0·5	1	—	—
26·75—	—	—	—	—	—	—	1	—	—	—	—	—
Totals	1	2	7	24·5	49	40	39·5	28·5	16·5	5·5	1·5	1

TABLE XXIV. Inter-tuberal.

TABLE XXV. Posterior sagittal.

Inter-spinous.

	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals	5·75—	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—
16·75—	—	1	0·5	—	—	—	—	—	—	—	—	1·5	—	—	1	—	0·5	—	—	—	—	—	—	—
17·25—	—	—	0·5	—	—	—	—	—	—	—	—	0·5	—	—	—	—	0·5	—	—	—	—	—	—	—
17·75	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—	—	—
18·25—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—	—	—
18·75—	—	2	—	—	1	—	—	—	—	—	—	3	1	1	—	1	—	—	—	—	—	—	—	—
19·25—	—	—	—	—	1	1	2	—	—	—	—	4	—	—	1	1	1	—	—	1	—	—	—	—
19·75—	1	1	1	1	2	2	—	—	—	—	—	8	2	2	1	2	1	—	—	—	—	—	—	—
20·25—	—	—	4	—	1	3	2	—	—	—	—	10	1	—	2	3·5	0·5	2	1	—	—	—	—	—
20·75—	—	2·5	2·5	2·5	3	1	1	—	—	0·5	0·5	13·5	—	—	3	4·5	4	1	1	—	—	—	—	—
21·25—	—	1	2·5	2·5	4	3	1	1	—	—	—	15	—	2	5·5	3	2	0·5	2	—	—	—	—	—
21·75—	—	3·5	6	6	13	3·5	3	1	—	—	—	36	—	6·5	10	7·5	8	2	1	—	—	—	—	1
22·25—	1·5	1·5	2	2	3	4·5	2	—	—	—	—	16·5	1	2	3·5	4	4	—	1	1	—	—	—	—
22·75—	0·5	1	7·5	6	3	3·5	1·5	—	1	—	—	24	—	3	5	5·5	5·5	3	1	1	—	—	—	—
23·25—	1·5	2	5	6	6	2	0·5	0·5	—	—	—	23·5	—	3·5	6	5·5	4	3·5	1	—	—	—	—	—
23·75—	2·5	3	2	3·5	5·5	6	4	2	—	—	—	28·5	—	4	7	7	6	3·5	—	1	—	—	—	—
24·25—	1	—	2	0·5	4·5	2	0·5	1	0·5	—	—	12	1	2	2	4	1	2	—	—	—	—	—	—
24·75—	—	1	0·5	1	2	3·5	1·5	—	—	—	—	9·5	—	2	—	1·5	3·5	0·5	1	—	—	1	—	—
25·25—	—	1	1·5	1	1	2	—	—	—	—	—	6·5	—	1	—	2	1	1·5	1	—	—	—	—	—
25·75—	—	—	—	0·5	—	—	—	—	—	1	—	1·5	—	—	—	0·5	—	1	—	—	—	—	—	—
26·25—	—	—	—	0·5	—	—	—	1	—	—	—	1·5	—	—	—	0·5	—	—	1	—	—	—	—	—
26·75—	1	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1	—	—	—	—	—	—	—
Totals	9	20·5	37·5	33	50	37	19	6·5	1·5	1·5	0·5	216	6	29	47	53	43·5	20·5	11	4	0	1	0	1

TABLE XXVI.

Inter-cristal.

Inter-spinous.	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	27·25—	27·75—	28·25—	28·75—	Totals
16·75—	1	—	0·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1·5
17·25—	—	—	0·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·5
17·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
18·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
18·75—	—	—	—	—	1	—	1	1	—	—	—	—	—	—	—	—	—	3
19·25—	—	—	—	—	1	1	1	1	—	—	—	—	—	—	—	—	—	4
19·75—	—	—	1	—	1	0·5	3·5	1	—	0·5	0·5	—	—	—	—	—	—	8
20·25—	—	—	—	—	1·5	0·25	2·25	1	2	3	—	—	—	—	—	—	—	10
20·75—	—	—	1	—	1	0·25	5·25	0·5	1·5	1	1·5	0·5	—	0·5	0·5	—	—	13·5
21·25—	—	—	—	—	1·5	—	0·5	1·5	5·5	1	2	1	1	1	—	—	—	15
21·75—	—	—	—	—	1	1	2	8·5	7·25	6·75	7	1·5	1	—	—	—	—	36
22·25—	—	—	—	—	—	—	1	2	3·75	2·25	3·25	2·25	2	—	—	—	—	16·5
22·75—	—	—	—	—	—	—	—	—	6	5	4·75	4·75	2·5	1	—	—	—	24
23·25—	—	—	—	—	—	—	—	—	3·5	4·5	6	1·75	5·75	—	2	—	—	23·5
23·75—	—	—	—	—	—	—	—	—	1	0·5	8·5	4·75	6·25	1·5	6	—	—	28·5
24·25—	—	—	—	—	—	—	—	—	—	—	0·75	1·25	7	1·5	1	0·5	—	12
24·75—	—	—	—	—	—	—	—	—	—	—	0·25	0·25	1·25	3·25	1	2·5	1	9·5
25·25—	—	—	—	—	—	—	—	—	—	—	—	—	1·75	2·75	1	1	—	6·5
25·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0·5	—	1·5
26·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	0·5	—	1·5
26·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1
Totals	1	0	3	0	8	3	16·5	16·5	30·5	24·5	34·5	18	28·5	12·5	12·5	6	1	216

TABLE XXVII. Diagonal conjugate.

Inter-cristal.	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	Totals
20·75—	—	—	1	—	—	—	—	—	—	—	—	—	1
21·25—	—	—	—	—	—	—	—	—	—	—	—	—	0
21·75—	—	1	—	1	—	1	—	—	—	—	—	—	3
22·25—	—	—	—	—	—	—	—	—	—	—	—	—	0
22·75—	—	1	—	1	—	2	2	2	—	—	—	—	8
23·25—	—	—	—	—	1·25	0·25	—	—	1·5	—	—	—	3
23·75—	—	—	—	2	6·25	3·25	2·5	1	1·5	—	—	—	16·5
24·25—	1	1	1	0·5	6	1	2	2	1	1	—	—	16·5
24·75—	—	1	1	4·5	5·5	8·5	2	2	5	—	—	1	30·5
25·25—	—	1	0·5	4	4	3·5	7·5	2	1	—	1	—	24·5
25·75—	—	1	1·5	6	7	4·5	6	4·5	2·5	1·5	—	—	34·5
26·25—	1	—	1	0·5	4	5	2	3	1	0·5	—	—	18
26·75—	—	—	—	4·5	3·5	4	5·5	6	2	1	2	—	28·5
27·25—	—	—	—	—	2	3	3	1	1	1	0·5	1	12·5
27·75—	—	—	—	—	1	4	3	2	2	—	0·5	—	12·5
28·25—	—	—	—	2	—	1	—	2	1	—	—	—	6
28·75—	—	—	—	—	—	—	—	—	—	—	1	—	1
Totals	2	6	6	26	40·5	41	35·5	27·5	19·5	5	5	2	216

TABLE XXVIII. Antero-posterior.

Inter-cristal.	Totals	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
20·75—	1	—	—	0·5	0·5	—	—	—	—	—	—	—	—
21·25—	0	—	—	—	—	—	—	—	—	—	—	—	—
21·75—	3	—	—	—	—	2	1	—	—	—	—	—	—
22·25—	0	—	—	—	—	—	—	—	—	—	—	—	—
22·75—	8	—	1	1	2	2	2	—	—	—	—	—	—
23·25—	3	—	—	—	—	0·5	1	1·5	—	—	—	—	—
23·75—	16·5	—	—	1	4·5	0·5	3·5	3	1	2	1	—	—
24·25—	16·5	—	—	1	2·5	5·5	3	3·25	0·25	1	—	—	—
24·75—	30·5	—	—	1	0·5	7·5	8	5·25	8·25	—	—	—	—
25·25—	24·5	—	—	1	4	5·75	2·75	6	3	1	1	—	—
25·75—	34·5	—	1	0·5	6	10·75	5·25	4	4·5	1	—	1	0·5
26·25—	18	—	—	—	1	4	4·25	6·25	1	1	—	—	0·5
26·75—	28·5	1	—	1	2·5	6·5	5·25	6·25	3·5	2·5	—	—	—
27·25—	12·5	—	—	—	—	1	2	1	3	3	2	0·5	—
27·75—	12·5	—	—	—	1	2	1	—	4	3	1·5	—	—
28·25—	6	—	—	—	—	1	1	3	—	1	—	—	—
28·75—	1	—	—	—	—	—	—	—	—	1	—	—	—
Totals	216	1	2	7	24·5	49	40	39·5	28·5	16·5	5·5	1·5	1

TABLE XXIX. Inter-tuberal.

Inter-cristal.	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
20·75—	—	1	—	—	—	—	—	—	—	—	—	1
21·25—	—	—	—	—	—	—	—	—	—	—	—	0
21·75—	—	—	2	—	—	1	—	—	—	—	—	3
22·25—	—	—	—	—	—	—	—	—	—	—	—	0
22·75—	1	1	2	1	1	1	1	—	—	—	—	8
23·25—	—	—	0·5	—	1·5	1	—	—	—	—	—	3
23·75—	—	4	1·5	0·5	5·5	2	3	—	—	—	—	16·5
24·25—	1	2	1	7	3·5	2	—	—	—	—	—	16·5
24·75—	—	—	5	6·5	9·5	4	4·5	1	—	—	—	30·5
25·25—	—	2	7·5	3	4	4·5	2·5	1	—	—	—	24·5
25·75—	2	4	6·5	5	6	9	1	—	0·5	0·25	0·25	34·5
26·25—	—	3	2	2·5	3·5	2·5	3	0·5	0·5	0·25	0·25	18
26·75—	4	2	4·5	3·5	7·5	3·5	1	2	0·5	—	—	28·5
27·25—	—	1·25	2·75	—	3	3·5	1	1	—	—	—	12·5
27·75—	—	0·25	1·25	2	4	2	1	1	—	1	—	12·5
28·25—	1	—	1	2	1	—	1	—	—	—	—	6
28·75—	—	—	—	—	—	1	—	—	—	—	—	1
Totals	9	20·5	37·5	33	50	37	19	6·5	1·5	1·5	0·5	216

TABLE XXX. Posterior sagittal.

Inter-cristal.	Totals	5·75—	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—
20·75—	1	—	—	1	—	—	—	—	—	—	—	—	—
21·25—	0	—	—	—	—	—	—	—	—	—	—	—	—
21·75—	3	—	—	1	1	1	—	—	—	—	—	—	—
22·25—	0	—	—	—	—	—	—	—	—	—	—	—	—
22·75—	8	2	—	2	3	—	1	—	—	—	—	—	—
23·25—	3	—	—	—	—	2	1	—	—	—	—	—	—
23·75—	16·5	—	2·5	4	3	3	1	2	1	—	—	—	—
24·25—	16·5	1	3·5	5·5	2·5	3	—	1	—	—	—	—	—
24·75—	30·5	—	3	7·5	10	3	4	1	2	—	—	—	—
25·25—	24·5	1·5	2	4	7·5	3·5	3	2	—	—	—	—	1
25·75—	34·5	0·5	10·5	8	7·25	6·25	1	—	1	—	—	—	—
26·25—	18	—	1·5	6	1·75	6·75	2	—	—	—	—	—	—
26·75—	28·5	1	3·5	6	8·5	5	2·5	2	—	—	—	—	—
27·25—	12·5	—	0·5	—	3	4·25	1·75	2	—	—	1	—	—
27·75—	12·5	—	1	2	2·5	2·75	3·25	1	—	—	—	—	—
28·25—	6	—	1	—	2	3	—	—	—	—	—	—	—
28·75—	1	—	—	—	1	—	—	—	—	—	—	—	—
Totals	216	6	29	47	53	43·5	20·5	11	4	0	1	0	1

TABLE XXXI. Antero-posterior.

Diagonal conjugate.	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	Totals
9·25—	—	—	—	—	—	2	—	—	—	—	—	—	2
9·75—	—	—	0·5	2·5	3	—	—	—	—	—	—	—	6
10·25—	—	—	1·5	0·5	1	1	1	1	—	—	—	—	6
10·75—	—	1	—	4	7·5	5·5	4·5	2·5	—	1	—	—	26
11·25—	—	—	1·5	5·5	9·5	7·5	8	5·5	1	0·5	0·5	1	40·5
11·75—	—	—	3	5·5	8	7·5	9	7	1	—	—	—	41
12·25—	1	1	—	3	11·5	4	6	3	4	1	1	—	35·5
12·75—	—	—	0·5	1·5	6	6·5	6	3	3	1	—	—	27·5
13·25—	—	—	—	2	1·5	5·5	2·5	5	2	1	—	—	19·5
13·75—	—	—	—	—	—	0·5	2·5	—	2	—	—	—	5
14·25—	—	—	—	—	1	—	—	0·5	2·5	1	—	—	5
14·75—	—	—	—	—	—	—	—	1	1	—	—	—	2
Totals	1	2	7	24·5	49	40	39·5	28·5	16·5	5·5	1·5	1	216

TABLE XXXII. Inter-tuberal.

Diagonal conjugate.	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—
9·25—	—	—	2	—	—	—	—	—	—	—	—
9·75—	—	1	1	3	1	—	—	—	—	—	—
10·25—	—	2	1	1	2	—	—	—	—	—	—
10·75—	3	4	4	4	5	3	2	1	—	—	—
11·25—	1	5	7	6	9·5	9	2	—	—	0·5	0·5
11·75—	—	5	6·5	7	7·5	8	5	1	—	1	—
12·25—	1	2	6·5	4	10	5	5	2	—	—	—
12·75—	3·5	—	6	4	6	5	1	2	—	—	—
13·25—	0·5	1	3	3	7	1	4	—	—	—	—
13·75—	—	—	—	—	—	3	—	0·5	1·5	—	—
14·25—	—	0·5	0·5	—	2	2	—	—	—	—	—
14·75—	—	—	—	1	—	1	—	—	—	—	—
Totals	9	20·5	37·5	33	50	37	19	6·5	1·5	1·5	0·5

TABLE XXXIII. Posterior sagittal.

Diagonal conjugate.	5·75—	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	Totals
9·25—	—	—	—	—	2	—	—	—	—	—	—	—	2
9·75—	—	2	2	1	—	1	—	—	—	—	—	—	6
10·25—	1	1	2	—	2	—	—	—	—	—	—	—	6
10·75—	1	7·5	4·5	5	3	3	1	1	—	—	—	—	26
11·25—	1	5·5	11·5	10	5·5	3	2	1	—	1	—	—	40·5
11·75—	2	5	13	11	7	2	—	1	—	—	—	—	41
12·25—	1	4	4	11	6·5	5	2	1	—	—	—	1	35·5
12·75—	—	3	3·5	9	9	2	1	—	—	—	—	—	27·5
13·25—	—	1	5·5	3	5	3	2	—	—	—	—	—	19·5
13·75—	—	—	—	1	2	—	2	—	—	—	—	—	5
14·25—	—	—	1	2	0·5	0·5	1	—	—	—	—	—	5
14·75—	—	—	—	—	1	1	—	—	—	—	—	—	2
Totals	6	29	47	53	43·5	20·5	11	4	0	1	0	1	216

TABLE XXXIV. Inter-tubereal.

Antero-posterior.	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
8·75—	—	—	—	—	1	—	—	—	—	—	—	1
9·25—	—	1	—	—	1	—	—	—	—	—	—	2
9·75—	—	2·5	1·5	2	1	—	—	—	—	—	—	7
10·25—	1	3·5	5·5	6	6	2	—	0·5	—	—	—	24·5
10·75—	2	3	11	6	16	6·5	2	2·5	—	—	—	49
11·25—	2	3	9	4	11	5·5	4	1·25	0·25	—	—	40
11·75—	2	7	6	6·5	6	6·5	4	0·25	1·25	—	—	39·5
12·25—	2	—	3	6·5	4·5	7·5	5	—	—	—	—	28·5
12·75—	—	—	1	2	2·5	8	1	1	—	1	—	16·5
13·25—	—	0·5	0·5	—	0·5	1	2	1	—	—	—	5·5
13·75—	—	—	—	—	0·5	—	1	—	—	—	—	1·5
14·25—	—	—	—	—	—	—	—	—	—	0·5	0·5	1
Totals	9	20·5	37·5	33	50	37	19	6·5	1·5	1·5	0·5	216

TABLE XXXV. Posterior sagittal.

Antero-posterior.	5·75—	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—
8·75—	—	—	—	1	—	—	—	—	—	—	—	—
9·25—	—	1	1	—	—	—	—	—	—	—	—	—
9·75—	2	1	3	0·5	—	0·5	—	—	—	—	—	—
10·25—	3	6	8·5	4·5	2	0·5	—	—	—	—	—	—
10·75—	1	13	13·5	12	5	3·5	1	—	—	—	—	—
11·25—	—	3	10·5	11·5	10	3	1	1	—	—	—	—
11·75—	—	2	7·5	9·5	14·5	3·5	1·5	1	—	—	—	—
12·25—	—	2	3	9·5	4·5	6	2·5	1	—	—	—	—
12·75—	—	—	—	4·5	6	3	3	—	—	—	—	—
13·25—	—	1	—	—	0·5	0·5	2	—	—	0·5	—	1
13·75—	—	—	—	—	—	—	—	1	—	0·5	—	—
14·25—	—	—	—	—	1	—	—	—	—	—	—	—
Totals	6	29	47	53	43·5	20·5	11	4	0	1	0	1

TABLE XXXVI. Inter-tuberal.

Posterior sagittal.	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
5·75—	1	1	1	—	1	2	—	—	—	—	—	6
6·25—	1	6	8·5	2·5	5	5	1	—	—	—	—	29
6·75—	3	9	4·5	11·5	13	2	2	2	—	—	—	47
7·25—	2	3	12·5	6·5	13	9	5	1·5	0·5	—	—	53
7·75—	1	1·25	9·75	6·5	11	9	3	—	1	0·5	0·5	43·5
8·25—	1	0·25	1·25	6	3	4	3	1	—	1	—	20·5
8·75—	—	—	—	—	2	6	1	2	—	—	—	11
9·25—	—	—	—	—	1	—	3	—	—	—	—	4
9·75—	—	—	—	—	—	—	—	—	—	—	—	0
10·25—	—	—	—	—	1	—	—	—	—	—	—	1
10·75—	—	—	—	—	—	—	—	—	—	—	—	0
11·25—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	9	20·5	37·5	33	50	37	19	6·5	1·5	1·5	0·5	216

APPENDIX III.

CORRELATION TABLES OF THE DIAMETERS WITH THE BONE MEASUREMENTS.

TABLE XXXVII. Obstetric conjugate.

Right hip—height.	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	Totals
15·75—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	1
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	2	1	—	—	—	—	—	—	—	—	—	—	3
17·25—	—	—	1	4	—	—	—	1	—	—	—	—	—	—	6
17·75—	1	1	1	5	3	3·5	6·5	2	—	—	—	—	—	—	23
18·25—	0·5	2·5	—	10	7	6	3·5	8	0·5	1	—	—	—	—	39
18·75—	—	—	2	5	5	13·5	5·5	7	4·5	1·5	—	—	—	—	44
19·25—	—	—	—	3	5	5	5·5	2·5	2	3	—	—	—	—	26
19·75—	—	—	—	1	3	8	6	9	6	5	—	2	—	—	40
20·25—	—	—	—	0·5	1·5	1	3	4·5	2·5	2	2	1	—	—	18
20·75—	—	—	—	—	—	2	—	3	2·5	0·5	—	1	—	—	9
21·25—	—	—	—	—	—	1	1	1	—	1	—	—	—	—	4
21·75—	—	—	—	—	—	—	1	—	—	1	—	—	—	1	3
Totals	1·5	3·5	7	29·5	24·5	40	32	38	18	15	2	4	0	1	216

TABLE XXXVIII. Transverse.

Right hip—height.	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
15·75—	1	—	—	—	—	—	—	—	—
16·25—	—	—	—	—	—	—	—	—	—
16·75—	—	—	—	1	1	1	—	—	—
17·25—	—	—	2	3	—	1	—	—	—
17·75—	—	—	1	7	4	6	3	1·5	0·5
18·25—	—	—	1·5	8·5	7	17	4	1	—
18·75—	1	—	1	4·5	8	20	6	3	0·5
19·25—	—	1	—	3	2	7·5	7·5	4	1
19·75—	—	—	1	1	7·5	17	7·5	5	1
20·25—	—	—	—	—	—	5·5	6·5	5	1
20·75—	—	—	—	—	—	3	2	3	1
21·25—	—	—	—	—	—	—	—	2	2
21·75—	—	—	—	1	—	1	1	—	—
Totals	2	1	6·5	29	29·5	79	37·5	24·5	7

TABLE XXXIX. Diagonal conjugate.

Right hip—height.	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	Totals
15·75—	—	—	1	—	—	—	—	—	—	—	—	—	1
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	1	1	1	—	—	—	—	—	—	—	—	3
17·25—	—	1	—	3	1	1	—	—	—	—	—	—	6
17·75—	1	2	2	2	8	2	4	2	—	—	—	—	23
18·25—	1	2	2	8	12·5	7·5	4	1	1	—	—	—	39
18·75—	—	—	—	8	10	10	6	6·5	3·5	—	—	—	44
19·25—	—	—	—	1	6	8·5	4·5	4	2	—	—	—	26
19·75—	—	—	—	3	2	8	7	9	7	2	2	—	40
20·25—	—	—	—	—	1	3	4	4	3	1	1	1	18
20·75—	—	—	—	—	—	1	3	1	2	2	—	—	9
21·25—	—	—	—	—	—	—	2	—	1	—	1	—	4
21·75—	—	—	—	—	—	—	1	—	—	—	1	1	3
Totals	2	6	6	26	40·5	41	35·5	27·5	19·5	5	5	2	216

TABLE XL. Inter-cristal.

Right hip—height.	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	27·25—	27·75—	28·25—	28·75—	Totals
15·75—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	—	—	1	—	—	1	1	—	—	—	—	—	—	—	—	3
17·25—	—	—	1	—	1	—	2	—	2	—	—	—	—	—	—	—	—	6
17·75—	—	—	—	—	2	—	5	5	2·5	1·5	3·5	0·5	1	1	—	1	—	23
18·25—	—	—	1	—	1	1·5	5·5	4	9·5	6	5·5	3	1	1	—	—	—	39
18·75—	—	—	—	—	2	1	2·5	3·5	5·5	4	11·5	6·5	6	0·5	1	—	—	44
19·25—	—	—	1	—	1	—	—	1	3	5·5	2·5	2	7	2	1	—	—	26
19·75—	—	—	—	—	—	—	1	1	3	7	7	3	9	2	5	2	—	40
20·25—	—	—	—	—	—	—	—	1	3	—	2·5	2·5	3·5	1·5	3	1	—	18
20·75—	—	—	—	—	—	0·5	0·5	—	1	0·5	2	0·5	1	2	—	1	—	9
21·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	1	1	4
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	2·5	0·5	—	—	3
Totals	1	0	3	0	8	3	16·5	16·5	30·5	24·5	34·5	18	28·5	12·5	12·5	6	1	216

TABLE XLI. Inter-spinous.

Right hip—height.	16·75—	17·25—	17·75—	18·25—	18·75—	19·25—	19·75—	20·25—	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	Totals
15·75—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	—	—	1	—	—	—	0·5	0·5	—	—	—	1	—	—	—	—	—	—	—	3
17·25—	—	—	—	—	2	—	—	—	2	1	—	—	1	—	—	—	—	—	—	—	—	6
17·75—	—	—	—	—	—	—	1	2·5	4·5	—	7	2	1	1	1	1	1	1	—	—	—	23
18·25—	0·5	0·5	—	—	—	2	4	—	2	2·5	6·5	9	6	3	1	1·5	0·5	—	—	—	—	39
18·75—	—	—	—	—	—	2	1	2	1	4	9·5	2·5	5·5	6·5	5	4	—	1	—	—	—	44
19·25—	—	—	—	—	—	—	2	3	—	2	4	0·5	3	3·5	6	—	1	1	—	—	—	26
19·75—	—	—	—	—	—	—	—	1	2	2	7	2	5	6·5	5·5	3·5	2·5	2	—	1	—	40
20·25—	—	—	—	—	—	—	—	1	—	2	2	0·5	1·5	1	6	1	0·5	0·5	1	—	1	18
20·75—	—	—	—	—	—	—	—	0·5	0·5	1	—	—	1	—	4	—	—	1	0·5	0·5	—	9
21·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	3	—	—	—	—	4
21·75—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1	1	—	—	—	—	3
Totals	1·5	0·5	0	0	3	4	8	10	13·5	15	36	16·5	24	23·5	28·5	12	9·5	6·5	1·5	1·5	1	216

TABLE XLII. Obstetric conjugate.

Left hip—height.	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	Totals
15·25—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	1
15·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	2
17·25—	—	—	3	4	—	—	—	—	—	—	—	—	—	—	7
17·75—	1	1	—	5	2	3·5	4·5	2	—	—	—	—	—	—	19
18·25—	—	1	—	11	7	6	3·5	8	0·5	1	—	—	—	—	38
18·75—	0·5	1·5	2	3	6	13·5	7·5	6	3	1	—	—	—	—	44
19·25—	—	—	—	4	4	6	5·5	4·5	3·5	3·5	—	—	—	—	31
19·75—	—	—	—	1	4	7	5	5	3	3	—	1	—	—	29
20·25—	—	—	—	0·5	1·5	2	3	8	5·5	4·5	2	2	—	—	29
20·75—	—	—	—	—	—	1	2	2·5	2·5	—	—	—	—	—	8
21·25—	—	—	—	—	—	1	1	2	—	2	—	1	—	—	7
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
22·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1
Totals	1·5	3·5	7	29·5	24·5	40	32	38	18	15	2	4	0	1	216

TABLE XLIII. Transverse.

Left hip—height.	Totals	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—
15·25—	1	1	—	—	—	—	—	—	—	—
15·75—	0	—	—	—	—	—	—	—	—	—
16·25—	0	—	—	—	—	—	—	—	—	—
16·75—	2	—	—	—	1	—	1	—	—	—
17·25—	7	—	—	1	3	1	2	—	—	—
17·75—	19	—	—	1	6	5	5	2	—	—
18·25—	38	1	—	2·5	7·5	6	16	2	2·5	0·5
18·75—	44	—	—	1	4·5	6	19	9	4	0·5
19·25—	31	—	1	—	3	4·5	11	6·5	4	1
19·75—	29	—	—	1	2	4·5	12	6·5	3	—
20·25—	29	—	—	—	—	2·5	10	8·5	5	3
20·75—	8	—	—	—	1	—	1	2	3	1
21·25—	7	—	—	—	—	—	2	1	3	1
21·75—	0	—	—	—	—	—	—	—	—	—
22·25—	1	—	—	—	1	—	—	—	—	—
Totals	216	2	1	6·5	29	29·5	79	37·5	24·5	7

TABLE XLIV. Diagonal conjugate.

Left hip—height.	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	Totals
15·25—	—	—	1	—	—	—	—	—	—	—	—	—	1
15·75—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	1	1	—	—	—	—	—	—	—	—	2
17·25—	—	3	—	3	1	—	—	—	—	—	—	—	7
17·75—	1	2	—	3	6	3	2	2	—	—	—	—	19
18·25—	—	—	4	7	13·5	5·5	5	2	1	—	—	—	38
18·75—	1	1	—	7	10	13	7	3	2	—	—	—	44
19·25—	—	—	—	2	7	8·5	3·5	5·5	4·5	—	—	—	31
19·75—	—	—	—	3	2	6	8	4	3	2	1	—	29
20·25—	—	—	—	—	1	4	3	10	7	1	2	1	29
20·75—	—	—	—	—	—	1	4	1	1	1	—	—	8
21·25—	—	—	—	—	—	—	3	—	1	1	2	—	7
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	0
22·25—	—	—	—	—	—	—	—	—	—	—	—	1	1
Totals	2	6	6	26	40·5	41	35·5	27·5	19·5	5	5	2	216

TABLE XLV.

Inter-cristal.

Left hip—height.	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	27·25—	27·75—	28·25—	28·75—	Totals
15·25—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
15·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	2
17·25—	—	—	1	—	2	—	1	1	2	—	—	—	—	—	—	—	—	7
17·75—	—	—	—	—	2	—	4	4	2·5	3·5	1	—	1	—	—	1	—	19
18·25—	—	—	1	—	1	1·5	8	4·5	6	3·5	7	2·5	2	1	—	—	—	38
18·75—	—	—	—	—	2	—	1	2	9	5·5	10·5	6·5	5	1·5	1	—	—	44
19·25—	—	—	—	—	1	1	1	2	4	2·5	4·5	3	8	2	2	—	—	31
19·75—	—	—	1	—	—	—	1	1	2	8	4	3	5	1	1	2	—	29
20·25—	—	—	—	—	—	0·5	0·5	1	3	1	6·5	2·5	6	3	4	1	—	29
20·75—	—	—	—	—	—	—	—	—	1	0·5	0·5	—	0·5	1·5	3	1	—	8
21·25—	—	—	—	—	—	—	—	—	—	—	0·5	0·5	1	1·5	1·5	1	1	7
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
22·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	1
Totals	1	0	3	0	8	3	16·5	16·5	30·5	24·5	34·5	18	28·5	12·5	12·5	6	1	216

TABLE XLVI.

Inter-spinous.

Left hip—height.	16·75—	17·25—	17·75—	18·25—	18·75—	19·25—	19·75—	20·25—	20·75—	21·25—	21·75—	22·25—	22·75—	23·25—	23·75—	24·25—	24·75—	25·25—	25·75—	26·25—	26·75—	Totals
15·25—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
15·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	—	—	1	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	2
17·25—	—	—	—	—	1	—	—	—	2·5	1·5	1	—	1	—	—	—	—	—	—	—	—	7
17·75—	—	—	—	—	1	1	1	0·5	2·5	1	5	2	1	1	1	1	1	—	—	—	—	19
18·25—	0·5	0·5	—	—	—	2	4	2	4	1·5	6·5	7	5	1	1	2·5	0·5	—	—	—	—	38
18·75—	—	—	—	—	—	1	1	2	1	3	8·5	4·5	7·5	8	2·5	3	—	2	—	—	—	44
19·25—	—	—	—	—	—	—	1	2	—	4	6	0·5	2	3	10	0·5	1	1	—	—	—	31
19·75—	—	—	—	—	—	—	1	2	2	—	5	2	4	6·5	1·5	1·5	2·5	1	—	—	—	29
20·25—	—	—	—	—	—	—	—	1·5	0·5	3	4	0·5	2·5	2	8·5	2·5	—	1	1	1	1	29
20·75—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	3	—	1·5	1·5	0·5	0·5	—	8
21·25—	—	—	—	—	—	—	—	—	1	—	—	—	1	1	1	1	2	—	—	—	—	7
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
22·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	1·5	0·5	0	0	3	4	8	10	13·5	15	36	16·5	24	23·5	28·5	12	9·5	6·5	1·5	1·5	1	216

TABLE XLVII. Right hip—breadth.

Obstetric conjugate.	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—	Totals
7·25—	—	—	—	—	1	0·5	—	—	—	—	—	—	1·5
7·75—	—	—	—	—	1	0·5	2	—	—	—	—	—	3·5
8·25—	1	—	—	3	1	1	1	—	—	—	—	—	7
8·75—	—	—	—	3·5	8	10·5	5	2·5	—	—	—	—	29·5
9·25—	—	—	—	1	5	8	3	5·5	1	1	—	—	24·5
9·75—	—	—	—	1·5	4·5	11	6	9	6	1	1	—	40
10·25—	—	—	—	1·5	2·5	5	7	6·5	6·5	2	1	—	32
10·75—	—	1	—	—	3	9	10	5·5	7·5	1	1	—	38
11·25—	—	—	—	—	—	3·5	5·5	2	4	3	—	—	18
11·75—	—	—	—	—	—	2·5	3	3	2·5	3	—	1	15
12·25—	—	—	—	—	—	—	—	—	—	2	—	—	2
12·75—	—	—	—	—	—	—	1	2	1	—	—	—	4
13·25—	—	—	—	—	—	—	—	—	—	—	—	—	0
13·75—	—	—	—	—	—	—	—	—	—	—	1	—	1
Totals	1	1	0	10·5	26	51·5	43·5	36	28·5	13	4	1	216

TABLE XLVIII. Left hip—breadth.

Obstetric conjugate.	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—
7·25—	—	—	—	—	—	1	0·5	—	—	—	—	—	—
7·75—	—	—	—	—	—	1	1·5	1	—	—	—	—	—
8·25—	1	—	—	0·5	2·5	1	2	—	—	—	—	—	—
8·75—	—	—	—	—	3	10·5	6·5	8	1	0·5	—	—	—
9·25—	—	—	—	—	—	5	7	6	4	2·5	—	—	—
9·75—	—	—	—	—	2·5	5	8·5	9·5	8·5	3·5	2·5	—	—
10·25—	—	—	—	—	2·5	0·5	6	7	8	5	3	—	—
10·75—	—	—	—	—	1	2	6	14	6	8	1	—	—
11·25—	—	—	—	—	—	0·5	2	4·5	4	6	1	—	—
11·75—	—	—	—	—	—	1·5	1	5	3	0·5	3	—	1
12·25—	—	—	—	—	—	—	—	—	—	—	2	—	—
12·75—	—	—	—	—	—	—	—	—	3	1	—	—	—
13·25—	—	—	—	—	—	—	—	—	—	—	—	—	—
13·75—	—	—	—	—	—	—	—	—	—	—	—	1	—
Totals	1	0	0	0·5	11·5	28	41	55	37·5	27	12·5	1	1

TABLE XLIX. Sacrum—height.

Obstetric conjugate.	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
7·25—	—	—	—	—	—	—	1·5	—	—	—	—	—	—	—	1·5
7·75—	—	—	—	1	—	1	0·5	1	—	—	—	—	—	—	3·5
8·25—	—	0·5	1	1·5	—	2	—	1	—	—	—	1	—	—	7
8·75—	—	—	—	—	4	5	9	5	1·5	2	1	2	—	—	29·5
9·25—	—	—	—	4	1	3	3	5	5·5	3	—	—	—	—	24·5
9·75—	—	—	1	2	3	5	9	6·5	4·5	3	2·5	1	1	1·5	40
10·25—	—	—	1	2	0·5	4	7	6	3·5	4	3·5	—	—	0·5	32
10·75—	1	—	—	2	1·5	2	3·5	8	5·5	8	4·5	2	—	—	38
11·25—	—	—	—	—	—	1	1·5	7·5	2	3·5	2·5	—	—	—	18
11·75—	—	—	1	—	1	—	—	2·5	1	5	1·5	2	1	—	15
12·25—	—	—	—	—	—	—	—	1	—	—	—	—	1	—	2
12·75—	—	—	—	—	—	—	1	1	—	1	1	—	—	—	4
13·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
13·75—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	1	0·5	4	12·5	11	23	36	44·5	23·5	30·5	16·5	8	3	2	216

TABLE L. Sacrum—breadth.

Obstetric conjugate.	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—
7·25—	—	—	—	—	—	1·5	—	—	—	—
7·75—	—	—	—	1	2	0·5	—	—	—	—
8·25—	1	—	—	3	2	1	—	—	—	—
8·75—	—	—	3	3	14	4	5	0·5	—	—
9·25—	—	—	0·5	5·5	7	5	6	0·5	—	—
9·75—	—	—	—	3	11·5	15	3·5	7	—	—
10·25—	—	—	—	1·5	5·5	14·5	7	2·5	—	1
10·75—	—	—	—	5·5	3	10·5	10·5	8·5	—	—
11·25—	—	—	—	—	2·5	6	4	4·5	1	—
11·75—	—	—	—	3	1·5	5·5	4	1	—	—
12·25—	—	—	—	—	—	—	—	2	—	—
12·75—	—	—	—	1	—	—	3	—	—	—
13·25—	—	—	—	—	—	—	—	—	—	—
13·75—	—	—	—	—	—	—	—	1	—	—
Totals	1	0	3·5	26·5	49	63·5	43	27·5	1	1

TABLE LI. Right hip—breadth.

Transverse.	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—	Totals
10·25—	1	—	—	—	—	1	—	—	—	—	—	—	2
10·75—	—	—	—	—	—	—	1	—	—	—	—	—	1
11·25—	—	—	—	0·5	2·5	2	0·5	1	—	—	—	—	6·5
11·75—	—	1	—	5	7·5	6	5·5	2	1	—	1	—	29
12·25—	—	—	—	1	4	9·5	7	3·5	3	1·5	—	—	29·5
12·75—	—	—	—	3	12	21	13	15·5	7·5	5	2	—	79
13·25—	—	—	—	—	—	9	11·5	6·5	7	3·5	—	—	37·5
13·75—	—	—	—	1	—	3	4·5	7	5	2	1	1	24·5
14·25—	—	—	—	—	—	—	0·5	0·5	5	1	—	—	7
Totals	1	1	0	10·5	26	51·5	43·5	36	28·5	13	4	1	216

TABLE LII. Left hip—breadth.

Transverse.	Totals	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—
10·25—	2	1	—	—	—	—	—	1	—	—	—	—	—	—
10·75—	1	—	—	—	—	—	—	—	1	—	—	—	—	—
11·25—	6·5	—	—	—	—	1	2	2	0·5	1	—	—	—	—
11·75—	29	—	—	—	0·5	5·5	5	8·5	7·5	—	1	—	1	—
12·25—	29·5	—	—	—	—	1	5·5	5	11·5	3·5	2·5	0·5	—	—
12·75—	79	—	—	—	—	4	13	15·5	17·5	17	6·5	5·5	—	—
13·25—	37·5	—	—	—	—	—	0·25	7·75	10	10·5	7	2	—	—
13·75—	24·5	—	—	—	—	—	2·25	1·25	6·5	3·5	6·5	3·5	—	1
14·25—	7	—	—	—	—	—	—	—	0·5	2	3·5	1	—	—
Totals	216	1	0	0	0·5	11·5	28	41	55	37·5	27	12·5	1	1

TABLE LIII. Sacrum—height.

Transverse.	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
10·25—	—	—	—	1	—	—	—	—	—	—	1	—	—	—	2
10·75—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1
11·25—	—	—	—	2	1	—	1	—	0·5	1	1	—	—	—	6·5
11·75—	—	—	0·5	1·5	2	6	4	8·5	1·5	4	—	1	—	—	29
12·25—	—	0·5	1·5	1·5	1	4	6·5	6·5	3	4	1	—	—	—	29·5
12·75—	1	—	—	4·5	2·5	7	17	22·5	5·5	9	5	3	1	1	79
13·25—	—	—	1	1	4·5	1	3·5	4·75	7·75	7	4	1	2	—	37·5
13·75—	—	—	1	1	—	4	3	1·25	4·25	3·5	3·5	2	—	1	24·5
14·25—	—	—	—	—	—	1	1	1	—	2	1	1	—	—	7
Totals	1	0·5	4	12·5	11	23	36	44·5	23·5	30·5	16·5	8	3	2	216

TABLE LIV. Sacrum—breadth.

Transverse.	Totals	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—
10·25—	2	1	—	—	—	—	—	1	—	—	—
10·75—	1	—	—	0·5	0·5	—	—	—	—	—	—
11·25—	6·5	—	—	1	3·5	2	—	—	—	—	—
11·75—	29	—	—	2	10	6	7	3	1	—	—
12·25—	29·5	—	—	—	4·5	11	8·75	4·25	1	—	—
12·75—	79	—	—	—	7	22·5	27·75	16	4·75	1	—
13·25—	37·5	—	—	—	1	4	14	7·75	10·75	—	—
13·75—	24·5	—	—	—	—	3·5	3	9	9	—	—
14·25—	7	—	—	—	—	—	3	2	1	—	1
Totals	216	1	0	3·5	26·5	49	63·5	43	27·5	1	1

APPENDIX IV.

CORRELATION TABLES OF THE BONE MEASUREMENTS WITH ONE ANOTHER.

TABLE LV. Right hip—breadth. TABLE LVI. Left hip—breadth.

Right hip—height.	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—	Totals	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—
15·75—	1	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—	—	—	—
16·75—	—	—	—	2	1	—	—	—	—	—	—	—	3	—	—	—	—	2	1	—	—	—	—	—	—	—
17·25—	—	—	—	3·5	2·5	—	—	—	—	—	—	—	6	—	—	—	0·5	2·5	3	—	—	—	—	—	—	—
17·75—	—	1	—	3	8	6	4	—	1	—	—	—	23	—	—	—	—	5	8	5	3	2	—	—	—	—
18·25—	—	—	—	—	8·5	16·5	11	3	—	—	—	—	39	—	—	—	—	1	5·5	16·5	14	2	—	—	—	—
18·75—	—	—	—	—	4	18	10·5	9·5	2	—	—	—	44	—	—	—	—	—	7·5	11·5	17	7	1	—	—	—
19·25—	—	—	—	2	2	8	6	6	2	—	—	—	26	—	—	—	—	1	3	6	9	6	1	—	—	—
19·75—	—	—	—	—	—	3	10	10	12	5	—	—	40	—	—	—	—	—	—	2	10	14	11	3	—	—
20·25—	—	—	—	—	—	—	1	5	4	7	1	—	18	—	—	—	—	—	—	—	1	3	8	6	—	—
20·75—	—	—	—	—	—	—	1	2·5	5·5	—	—	—	9	—	—	—	—	—	—	—	1	2·5	5	0·5	—	—
21·25—	—	—	—	—	—	—	—	—	1	1	1	1	4	—	—	—	—	—	—	—	—	—	1	2	—	1
21·75—	—	—	—	—	—	—	—	—	1	—	2	—	3	—	—	—	—	—	—	—	—	1	—	1	1	—
Totals	1	1	0	10·5	26	51·5	43·5	36	28·5	13	4	1	216	1	0	0	0·5	11·5	28	41	55	37·5	27	12·5	1	1

TABLE LVII. Left hip—height. TABLE LVIII. Sacrum—breadth.

Right hip—height.	15·25—	15·75—	16·25—	16·75—	17·25—	17·75—	18·25—	18·75—	19·25—	19·75—	20·25—	20·75—	21·25—	21·75—	22·25—	Totals	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—
15·75—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—
16·75—	—	—	—	2	1	—	—	—	—	—	—	—	—	—	—	3	—	—	—	1	2	—	—	—	—	—
17·25—	—	—	—	—	5	1	—	—	—	—	—	—	—	—	—	6	—	—	2	3	1	—	—	—	—	—
17·75—	—	—	—	—	1	14	7	1	—	—	—	—	—	—	—	23	—	—	—	4	10	7	2	—	—	—
18·25—	—	—	—	—	—	4	28	7	—	—	—	—	—	—	—	39	—	—	1	10	13	12	—	3	—	—
18·75—	—	—	—	—	—	—	3	34	6	1	—	—	—	—	—	44	—	—	—	4	15	11·5	11·5	2	—	—
19·25—	—	—	—	—	—	—	—	2	21	3	—	—	—	—	—	26	—	—	0·5	1·5	3	11	6	4	—	—
19·75—	—	—	—	—	—	—	—	—	4	25	10	1	—	—	—	40	—	—	—	2	4	13	15	6	—	—
20·25—	—	—	—	—	—	—	—	—	—	—	16	2	—	—	—	18	—	—	—	1	1	3	4·5	7·5	1	—
20·75—	—	—	—	—	—	—	—	—	—	—	3	4	2	—	—	9	—	—	—	—	—	3	3	3	—	—
21·25—	—	—	—	—	—	—	—	—	—	—	—	1	3	—	—	4	—	—	—	—	—	1	1	1	—	1
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	2	—	1	3	—	—	—	—	—	2	—	1	—	—
Totals	1	0	0	2	7	19	38	44	31	29	29	8	7	0	1	216	1	0	3·5	26·5	49	63·5	43	27·5	1	1

TABLE LIX.

Sacrum—height.

Right hip—height.	*6·25—*	*6·75—*	*7·25—*	*7·75—*	*8·25—*	*8·75—*	*9·25—*	*9·75—*	*10·25—*	*10·75—*	*11·25—*	*11·75—*	*12·25—*	*12·75—*	Totals
15·75—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	1
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	0·5	0·5	—	—	1	—	1	—	—	—	—	—	—	3
17·25—	—	—	0·5	1·5	2	—	1	1	—	—	—	—	—	—	6
17·75—	—	—	—	1	—	8	5	5	1	1	2	—	—	—	23
18·25—	1	—	—	3	3	6	11	4	4	3	2	1	—	1	39
18·75—	—	—	—	2	3	4	7	11·5	4·5	6	4	2	—	—	44
19·25—	—	—	—	2	2	—	5	4	5	7	—	—	—	1	26
19·75—	—	—	2	2	1	2	5	9	6	4	4	3	2	—	40
20·25—	—	—	—	—	—	1	1	6	3	4	2	—	1	—	18
20·75—	—	—	1	—	—	1	—	2	—	2·5	1·5	1	—	—	9
21·25—	—	—	—	—	—	—	1	—	—	1	1	1	—	—	4
21·75—	—	—	—	—	—	—	—	1	—	2	—	—	—	—	3
Totals	1	0·5	4	12·5	11	23	36	44·5	23·5	30·5	16·5	8	3	2	216

TABLE LX.

Left hip—height.

Right hip—breadth.	*15·25—*	*15·75—*	*16·25—*	*16·75—*	*17·25—*	*17·75—*	*18·25—*	*18·75—*	*19·25—*	*19·75—*	*20·25—*	*20·75—*	*21·25—*	*21·75—*	*22·25—*	Totals
11·25—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
11·75—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	1
12·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
12·75—	—	—	—	1	4·5	2	1	1	1	—	—	—	—	—	—	10·5
13·25—	—	—	—	1	2·5	7	8·5	5	2	—	—	—	—	—	—	26
13·75—	—	—	—	—	—	6	15·5	15	10	5	—	—	—	—	—	51·5
14·25—	—	—	—	—	—	1	11	12·5	8	7	4	—	—	—	—	43·5
14·75—	—	—	—	—	—	1	2	8·5	7	7	7	2·5	1	—	—	36
15·25—	—	—	—	—	—	1	—	2	2	7	10	3·5	3	—	—	28·5
15·75—	—	—	—	—	—	—	—	—	1	3	7	2	—	—	—	13
16·25—	—	—	—	—	—	—	—	—	—	—	1	—	2	—	1	4
16·75—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1
Totals	1	0	0	2	7	19	38	44	31	29	29	8	7	0	1	216

TABLE LXI. Left hip—breadth.

Right hip—breadth	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—	Totals
11·25—	1	—	—	—	—	—	—	—	—	—	—	—	—	1
11·75—	—	—	—	—	—	—	1	—	—	—	—	—	—	1
12·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
12·75—	—	—	—	0·5	8·5	1·5	—	—	—	—	—	—	—	10·5
13·25—	—	—	—	—	3	17·75	4·25	1	—	—	—	—	—	26
13·75—	—	—	—	—	—	8·25	32·25	8	3	—	—	—	—	51·5
14·25—	—	—	—	—	—	0·5	2·5	35·5	5	—	—	—	—	43·5
14·75—	—	—	—	—	—	—	1	9·5	22	3·5	—	—	—	36
15·25—	—	—	—	—	—	—	—	1	7·5	19·5	0·5	—	—	28·5
15·75—	—	—	—	—	—	—	—	—	—	3	10	—	—	13
16·25—	—	—	—	—	—	—	—	—	—	1	2	1	—	4
16·75—	—	—	—	—	—	—	—	—	—	—	—	—	1	1
Totals	1	0	0	0·5	11·5	28	41	55	37·5	27	12·5	1	1	216

TABLE LXII. Sacrum—breadth.

Right hip—breadth	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—
11·25—	1	—	—	—	—	—	—	—	—	—
11·75—	—	—	—	1	—	—	—	—	—	—
12·25—	—	—	—	—	—	—	—	—	—	—
12·75—	—	—	1·5	3	4	2	—	—	—	—
13·25—	—	—	1·5	5	15·5	2	1	1	—	—
13·75—	—	—	—	8	16·5	19	6	2	—	—
14·25—	—	—	0·5	5·5	5	14	12	6·5	—	—
14·75—	—	—	—	2	6	11	10	7	—	—
15·25—	—	—	—	—	1	13·5	8	6	—	—
15·75—	—	—	—	2	1	—	4	4	1	1
16·25—	—	—	—	—	—	2	1	1	—	—
16·75—	—	—	—	—	—	—	1	—	—	—
Totals	1	0	3·5	26·5	49	63·5	43	27·5	1	1

TABLE LXIII. Sacrum—height.

Right hip—breadth.	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
11·25—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	1
11·75—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1
12·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
12·75—	—	0·5	1	0·5	1·5	2	3	1	—	1	—	—	—	—	10·5
13·25—	—	—	—	3	2·5	6·5	4	6	2	—	2	—	—	—	26
13·75—	1	—	—	4	1	3·5	12	13	4	9	3	1	—	—	51·5
14·25—	—	—	1	1	4	4	8	7·5	7·5	4·5	2	3	—	1	43·5
14·75—	—	—	1	2	—	4	6	6	3	7·5	4·5	—	1	1	36
15·25—	—	—	1	—	1	3	2	5	5	5·5	2	3	1	—	28·5
15·75—	—	—	—	1	1	—	—	4	1	2	3	—	1	—	13
16·25—	—	—	—	—	—	—	1	2	—	1	—	—	—	—	4
16·75—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1
Totals	1	0·5	4	12·5	11	23	36	44·5	23·5	30·5	16·5	8	3	2	216

TABLE LXIV. Left hip—breadth. TABLE LXV. Sacrum—breadth.

Left hip—height.	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—	13·75—	14·25—	14·75—	15·25—	15·75—	16·25—	16·75—	Totals	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—
15·25—	1	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—
15·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—
16·75—	—	—	—	—	1	1	—	—	—	—	—	—	—	2	—	—	—	1	1	—	—	—	—	—
17·25—	—	—	—	0·5	3·5	3	—	—	—	—	—	—	—	7	—	—	2	3	2	—	—	—	—	—
17·75—	—	—	—	—	3	7	7	1	1	—	—	—	—	19	—	—	—	6	6	6	1	—	—	—
18·25—	—	—	—	—	3	6·5	11·5	16	1	—	—	—	—	38	—	—	1	9	14	10	2	2	—	—
18·75—	—	—	—	—	—	5·5	13·5	15	9	1	—	—	—	44	—	—	—	3	15	14·5	8·5	3	—	—
19·25—	—	—	—	—	1	4	6	11	7	1	1	—	—	31	—	—	0·5	0·5	6	10	9	5	—	—
19·75—	—	—	—	—	—	1	3	7	9	7	2	—	—	29	—	—	—	3	4	11	8	3	—	—
20·25—	—	—	—	—	—	—	—	5	7·5	11·5	5	—	—	29	—	—	—	1	1	7	11·5	7·5	1	—
20·75—	—	—	—	—	—	—	—	—	1	4·5	2·5	—	—	8	—	—	—	—	—	2	1	4	—	1
21·25—	—	—	—	—	—	—	—	—	2	2	2	—	1	7	—	—	—	—	—	3	2	2	—	—
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—
22·25—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	—	—	1	—	—
Totals	1	0	0	0·5	11·5	28	41	55	37·5	27	12·5	1	1	216	1	0	3·5	26·5	49	63·5	43	27·5	1	1

TABLE LXVI. Sacrum—height.

Left hip—height.	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
15·25—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	1
15·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
16·75—	—	—	—	—	—	1	—	1	—	—	—	—	—	—	2
17·25—	—	0·5	1	0·5	2	1	1	1	—	—	—	—	—	—	7
17·75—	—	—	—	2	—	6	5	3	1	1	—	1	—	—	19
18·25—	1	—	—	3	4	7	11	4	2	2	3	—	—	1	38
18·75—	—	—	—	1	2	3	7	11·5	5·5	8	4	2	—	—	44
19·25—	—	—	—	4	2	1	6	5	5	6	1	—	—	1	31
19·75—	—	—	2	1	1	2	4	7	5	2	2	2	1	—	29
20·25—	—	—	—	—	—	1	1	11	5	4	3	2	2	—	29
20·75—	—	—	1	—	—	1	—	—	—	4	2	—	—	—	8
21·25—	—	—	—	—	—	—	1	1	—	2·5	1·5	1	—	—	7
21·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
22·25—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	1	0·5	4	12·5	11	23	36	44·5	23·5	30·5	16·5	8	3	2	216

TABLE LXVII. Sacrum—height.

Left hip—breadth.	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
10·75—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	1
11·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
11·75—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
12·25—	—	—	0·25	0·25	—	—	—	—	—	—	—	—	—	—	0·5
12·75—	—	0·5	0·75	0·25	1	2	3	3	—	—	1	—	—	—	11·5
13·25—	—	—	—	3	2	6·5	4	6·25	1·25	3	2	—	—	—	28
13·75—	—	—	—	4	1	3·5	8	10·25	4·25	6	2	2	—	—	41
14·25—	1	—	1	2	5	5	13	8	7	8	1	2	1	1	55
14·75—	—	—	1	1	1	4	6	7	4	5·5	4·5	2·5	—	1	37·5
15·25—	—	—	0·5	—	1	2	1	6	6	5	4	0·5	1	—	27
15·75—	—	—	0·5	1	—	—	1	4	1	2	2	—	1	—	12·5
16·25—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1
16·75—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1
Totals	1	0·5	4	12·5	11	23	36	44·5	23·5	30·5	16·5	8	3	2	216

TABLE LXVIII. Sacrum—breadth.

Left hip—breadth.	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	13·25—
10·75—	1	—	—	—	—	—	—	—	—	—
11·25—	—	—	—	—	—	—	—	—	—	—
11·75—	—	—	—	—	—	—	—	—	—	—
12·25—	—	—	—	0·5	—	—	—	—	—	—
12·75—	—	—	1	2·5	7	1	—	—	—	—
13·25—	—	—	1	6	16	4	—	1	—	—
13·75—	—	—	1	8	9	17	4	2	—	—
14·25—	—	—	0·5	6·5	10	15·5	15·5	7	—	—
14·75—	—	—	—	1	5	16	11·5	4	—	—
15·25—	—	—	—	—	2	8	7·5	9·5	—	—
15·75—	—	—	—	2	—	2	3·5	3	1	1
16·25—	—	—	—	—	—	—	—	1	—	—
16·75—	—	—	—	—	—	—	1	—	—	—
Totals	1	0	3·5	26·5	49	63·5	43	27·5	1	1

TABLE LXIX. Sacrum—height.

Sacrum—breadth.	6·25—	6·75—	7·25—	7·75—	8·25—	8·75—	9·25—	9·75—	10·25—	10·75—	11·25—	11·75—	12·25—	12·75—	Totals
8·75—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	1
9·25—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
9·75—	—	—	—	—	2	1	—	—	0·5	—	—	—	—	—	3·5
10·25—	—	—	0·5	4·5	1	5	5	3	4·5	2	—	1	—	—	26·5
10·75—	1	0·5	0·5	2	3	7	11	13·5	1·5	4	4	—	—	1	49
11·25—	—	—	1	2	1	6	13·5	14	6	13	3	2	1	1	63·5
11·75—	—	—	2	2	2	4	5·5	10·5	4	4	4	5	—	—	43
12·25—	—	—	—	1	2	—	1	2·5	7	6·5	5·5	—	2	—	27·5
12·75—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1
13·25—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	1	0·5	4	12·5	11	23	36	44·5	23·5	30·5	16·5	8	3	2	216

MISCELLANEA.

I. The Statistical Study of Dietaries.

By KARL PEARSON, F.R.S.

There is not the least doubt that a proper statistical study of diet would be of very great value. Such a study is of first class importance when we turn to the accurate discussion of working-class budgets. A statistical study of dietaries should consist of two parts. In the first section an analysis should be made of the food consumed and of the physiological value of its constituents for, say, a hundred working-class families, noting their income, occupation, rent, size of house and other details. In the next place an anthropometric, and where possible medical survey should be made of the families, their numbers, ages, statures, weights and general healths should be recorded, special pathological or disease conditions noted, and some record made of the habits of house-mother and father. Only when this has been fully, accurately and extensively enough done will it be possible to draw conclusions of real scientific value, and to give advice to the working-classes as to the best expenditure of income on foods.

A recent *Report upon a Study of the Diet of the Labouring Classes in the City of Glasgow* by Miss Dorothy E. Lindsay, issued by the authority of the Corporation and with a preface by Professor D. Noel Paton, M.D., has been widely reviewed, and various suggestions made therein have been cited as demonstrated facts without any critical examination of the data upon which they are based. Thus we have been told that a return to the national dish of porridge and milk is an urgent need of the Scottish working-classes, and Miss Lindsay's data have been cited as showing the inadequacy of the dietary of families with regular wages under 20*s.* per week. Miss Lindsay obtained dietaries of 60 working-class Glasgow families, and these dietaries show a very large amount of work and a considerable fulness of record. To the extent of 60 families she has certainly fulfilled the first requisite of an adequate enquiry as to the effectiveness of diet. Her families contained 400 individuals, of whom 246 were under 16 years of age. If we are to judge of the relative value of the various dietaries, this can only be done by discussing their effects on the 400 persons who partook of them. In order to obtain real knowledge from Miss Lindsay's results, this population ought to have been weighed and measured. Without this the dietaries seem to us of exceedingly small worth. Actually Miss Lindsay did ascertain the height and weight of a few children, put together in Appendix IV. as *Physical Condition of Children*.

This contains the heights of 10 girls and 7 boys, and the weights of 20 girls and 16 boys. These children are scattered over all ages from 5 to 13 and over the whole of the Groups A to L into which Miss Lindsay divides her families. If we only class her groups into three there is not even the weight of one boy and one girl for each year of age upon which to base any measure of the value of the dietaries! As an illustration of this let us take the important Group E of families with regular wages under 20*s.* Miss Lindsay, p. 18, provides the dietaries of five such

families containing 29 individuals, of whom 19 were children under 16. She writes of this Group: "In this section, which embraces those who may properly be called poor, not one diet reaches the minimum energy value of 3000 calories" (p. 18). Criticising individual diets she says: "In LI. the use of a greater proportion of the cheaper vegetable foods, e.g. potatoes and haricot beans, gave a better energy value for the same cost....In XXV. the animal protein is considerably in excess of the vegetable protein (see Appendix III.). As the former is more expensive, this diet could be improved and a higher protein content obtained by reversing the proportion of animal and vegetable protein. In view of the fact that the fat is low, necessarily so because of its cost, the carbohydrate intake should have been greater."

Dr Chalmers in his recent paper before the Royal Society of Medicine, "The House as a Contributory Factor in the Deathrate*," notes that Miss Lindsay's observations bear so directly on the inadequacy of the dietary of persons in the smaller houses that he quotes her sample dietaries in families with regular wages under 20*s.* per week and concludes with Miss Lindsay's remark that "the children are nearly all small and light in weight." It is not clear whether Miss Lindsay intended her remark to apply to Family XXV. only, or as Dr Chalmers supposes to the whole of Group E. Dr Chalmers' interpretation has been widely taken to be the true one although which is the correct interpretation matters little: *For of the individuals in this group not a single boy's weight is provided by Miss Lindsay, and of the girls in the group only two were measured.* These two are certainly below weight; they do not belong to Family XXV. but one to Family XVII. who shows developed rickets, and whose father has been a heavy drinker, and the other to Family LII. where the father suffers from phthisis and is in receipt of parish relief. It would therefore appear that Miss Lindsay's remark must apply to the group in general. But the whole demonstration of the effect of the dietaries is thus seen to turn on the weight of two individual girls aged 9 and 5 years respectively, one of whom has rickets†, and both of whom have degenerate fathers!

In order to test what relation, if any, the income of family has to weight of child, the correlation coefficients of weight (w) and income (i) for constant age have been found for both girls and boys for the whole of the 20 and 16 cases respectively provided by Miss Lindsay. We have‡:

Weight of Girls and Income for constant Age

$$_a r_{iw} = \cdot 03 \pm \cdot 15,$$

Weight of Boys and Income for constant Age

$$_a r_{iw} = \cdot 22 \pm \cdot 16.$$

In neither case is there any significance having regard to the probable error. Whether the weight of the child is or is not related to the parental income cannot possibly be settled on the data provided by Miss Lindsay, and, whether she possesses it or not, she publishes in her paper no material on which it would be possible to tell the effect of the dietaries of families with incomes under 20*s.* on the size and weight of the children.

But there are a great many other question-begging conclusions in Miss Lindsay's memoir. She starts apparently with the opinion that the physiological construction of the food is of no importance, and that a "calorie" whether obtained from peas or eggs, beans or beef is of equal value. But surely this is the very sort of question that an investigation of this kind should answer, not assume? Miss Lindsay criticises—at times severely—the housewife who spends her

* *Proceedings of the Royal Society of Medicine*, Vol. VI. (Section *Epidemiology and State Medicine*), pp. 155—181.

† Rickets can hardly be due to lowness of "calories" for some of the families with the highest calories (II. with 4003, and IV. with 3882) and high wages (IV. with 50*s.* 8*d.*, V. with 39*s.*, XXXIX. with 41*s.*) are rickety.

‡ I have most heartily to thank my colleague Miss Ethel M. Elderton for the values of most of the correlation coefficients of this note.

money on animal rather than on vegetable food, "Protein from vegetable sources is from one and a half to two times as much as that from animal sources. In one or two cases it is nearly three times as great, and in these diets the energy value received is remarkably high for the money laid out....In some of the studies this proportion of animal to vegetable protein is reversed....All these diets show the disadvantages of this large use of the animal food. The energy value is low and the cost is high. *How are these diets to be improved?* Study XIX. (Group G) seems to give the answer. In this family 'porridge was eaten twice a day.' The energy procured per penny spent was the second highest recorded, 619 calories, while it is noted that the children were strong, healthy and well-grown" (p. 28). And again :

"A porridge and milk diet contains the food principles in correct proportion. The protein-rich animal foods, flesh, fish, eggs, etc., are all too expensive for the labouring classes, and any increase in their amount in the diet is impracticable. But cheese and the cheap protein-rich vegetable foods, oatmeal, peas, beans, etc., should be more freely used" (p. 29). An examination of Study XIX. thus commended shows that 5·07 shillings were spent on animal, 8·75 shillings on vegetable food, or the expenditures were as 1 to 1·72, while in the population at large 698·77 shillings were spent on animal, 487·28 shillings on vegetable food, or the ratio was 1 to 0·72. This abnormal ratio is obtained in the family in question by taking no eggs, hardly any butter, and instead of milk, buttermilk, skimmilk and condensed milk; there is further no fish; the increased vegetable expenditure is on meal and potatoes. Now it would appear that before a statement is made that this is the right course to recommend, it is highly desirable to ascertain whether the money spent on vegetable or animal food is the more highly correlated with physical fitness in the individual. This can only be done with Miss Lindsay's material for the weights of the 20 girls and the 16 boys. Accordingly for these sparse data the correlations between A the expenditure per individual on animal food and V the expenditure per individual on vegetable food with weight of child for constant age were determined. These gave :

	Girls	Boys
${}_ar_{Aw}$ ...	$\cdot 16 \pm \cdot 15$	$\cdot 07 \pm \cdot 17$,
${}_ar_{Vw}$ ...	$\cdot 07 \pm \cdot 15$	$\cdot 30 \pm \cdot 15$.

If any stress could be laid on these results, we should have to conclude that money spent on animal food would be best in the case of girls, and money spent on vegetable food in the case of boys. But an examination of the probable errors shows that no weight whatever can be laid on the results. In fact that Miss Lindsay's data are wholly insufficient to answer the question of whether money spent on vegetable or animal food is the better. In fact if we take the correlation between weight w and F pence per individual spent on food we find :

	Girls	Boys
${}_ar_{Fw}$ ...	$\cdot 15 \pm \cdot 15$	$\cdot 22 \pm \cdot 16$,

both of which are low correlations and neither definitely significant having regard to their probable error.

Next the relation of weight in the children to the number of calories in animal food (C_A), the number of calories in vegetable food (C_V) and to the total number of calories in both (C_F) was investigated.

The following values were found :

	Girls	Boys
${}_ar_{C_Aw}$ ...	$\cdot 16 \pm \cdot 15$	$\cdot 51 \pm \cdot 12$,
${}_ar_{C_Vw}$ ...	$\cdot 25 \pm \cdot 14$	$\cdot 30 \pm \cdot 15$,
${}_ar_{C_Fw}$ ...	$\cdot 20 \pm \cdot 14$	$\cdot 41 \pm \cdot 14$.

If these results could be trusted at all, we should have to assert: (i) that the number of calories in the food is twice as influential in the case of the boy as in that of the girl, and

(ii) that the number of calories in the animal food is more influential than the number of calories in the vegetable food for the case of the boy, and the reverse is true for the case of the girl*. As a matter of fact the size of the probable errors shows us merely that it is very unsafe to draw any conclusions at all. On the basis of the scanty data provided by Miss Lindsay for the physical fitness of the individuals subjected to the various dietaries, no legitimate conclusions whatever can be drawn as to how wages or diet affect the individual; still less can it be asserted that a return to porridge and milk, and an avoidance of the purchase of flesh, fish and eggs would save the situation.

Indeed as far as any stress whatever can be laid on her slender material the conclusion to be drawn is entirely the other way, i.e. the greater the consumption of animal food relative to vegetable food the heavier will be the child. To test this the ratio of money spent on vegetable food (V) to the money spent on animal food (A) per individual was correlated with the weight of the child for constant age, i.e. the partial correlation ${}_a r_{w,V/A}$ was found. Further as the money spent on food might not have been spent to the best advantage the ratio of the number of calories in the vegetable food C_V to the number in the animal food C_A was also correlated with the weight for constant age, i.e. the partial correlation coefficient ${}_a r_{C_V/C_A,w}$ was found. The following values were determined:

		Girls	Boys
${}_a r_{V/A,w}$	...	$-\cdot 32 \pm \cdot 14$	$-\cdot 25 \pm \cdot 16$,
${}_a r_{C_V/C_A,w}$	...	$-\cdot 12 \pm \cdot 15$	$-\cdot 23 \pm \cdot 16$.

Now the correlations are not very large as compared with their probable errors, but they have *all one sign*, and this is negative. In other words the increase of expenditure on vegetable food relative to the expenditure on animal food, or the increase in the number of calories obtained from vegetable food as compared with the number obtained from animal food is associated in every case with *decreased* weight of the children. Thus we see that as far as any conclusion whatever can be drawn from Miss Lindsay's data, it is directly opposed to her statement that it is better to spend money on oatmeal, peas or beans than on flesh, fish, eggs, etc. She has started with the dogma that a calorie is of equal value whatever its origin, and not stayed to investigate whether it was even justified by her own material. Meanwhile her statement as to oatmeal has gone out as if it were a statistically demonstrated fact, whereas the only conclusion which would be justified in any measure by her data is that the proportion of animal to vegetable food should be kept as high as possible, if weight be taken as a test—although but a partial one—of the efficiency of a diet for a growing child.

The elaborate dietaries carefully worked out by Miss Lindsay are of very small service indeed, because they have not been accompanied by any adequate anthropometric record of the families thus dieted. The present reviewer wrote at once to Miss Lindsay and later to Professor Noel Paton in the hope that it might still be possible to save the situation by some attempt to weigh at least the majority of the members of the families whose diets were recorded. This appears, however, to be no longer possible†, and the result is the not unfamiliar one—an elaborate piece of investigation has been carried out and practically no safe conclusions can be drawn from it. Those concerned in it have not studied beforehand what would be needful statistically to establish an inference, and the sections of the memoir (pp. 28—32) on "The possibility of Improving these Diets" and on "The Relationship of these Diets to Health"—which have been most widely quoted in the popular press—may express either correct or incorrect views; the data provided are wholly insufficient to justify any conclusions at all.

* A little consideration will show that this does not confute the result reached later, because in ${}_a r_{C_A w}$ and ${}_a r_{C_V w}$ no attention is paid to the *relative* amounts of animal and vegetable food.

† A final appeal to Dr Chalmers, since this paper was in type, has resulted in his most kindly attempting all that was possible under the circumstances. Some 64 children of from 5 to 14 years of age have been found and measured, and I hope in the next number of this Journal to revise the above numerical constants on this larger, but still sadly slender material.

II. Note on the Surface of Constant Association.

By KARL PEARSON, F.R.S.

In the recent memoir by Dr Heron and myself a reference was made to the future publication of this note on the surface of constant association, *Biometrika*, Vol. IX, p. 315. The integral equation to the general surface was given, and also the particular form it took, still very complicated, for the simple case of total or 'marginal' frequencies being normal distributions. It was considered worth while to fully analyse one case of such a surface, namely that of $Q=0{\cdot}6$ for Gaussian marginal frequencies. The numerical calculations were carried out by Miss Julia Bell, M.A., and from her ordinates of the sections the sections were plotted by Mr H. E. Soper, M.A. with the aid of a Coradi coordinatograph. By interpolation when needful Mr Soper constructed the isoplethes of this surface of constant association, and also an excellent card model, for comparison with the model of a normal surface. The chief features of this surface were referred to in the paper just cited. Although the marginal frequencies are symmetrical, the cross sections are skew-curves, the skewness increasing from zero for the central section to ·16 for the y-array when $x=1{\cdot}5\sigma_x$ and to ·20 for the y-array when $x=3{\cdot}5\sigma_x$.

Diagram I gives the series of sections on one side of the mid-section $x=0$ up to $x=3{\cdot}5\sigma_x$ by intervals of $0{\cdot}25\sigma_x$; the same sections are repeated in inverse order on the other side of the central section $x=0$. It will be seen at once from the indicated means how skew the sections are.

Diagram II gives the isoplethes or contour-lines of equal frequency. They are approximately but not accurately ellipses with common principal axial directions, but they are very far indeed from being *similar* ellipses. In the contour corresponding to $z=\frac{9}{10}$ of its maximum value, the major axis is considerably more than twice the minor axis of the oval; in the contour corresponding to $z=\frac{1}{10}$, the major axis is very much less than twice the minor; the ovals tend indeed to less and less ellipticity. On this diagram are also plotted the regression lines of means and of modes. It will be seen that they tend to become parallel to the axis of x, or the regression tends to become zero. There is little doubt, that quite apart from normality of the marginal frequencies, any symmetrical marginal frequency would lead to like results, i.e. the isoplethes would not be similar curves, and the regression would be skew, and tend to asymptote to the horizontal. Thus the constancy of association would depend for its application on the existence of material in which the variates would be intimately related near their mean values and cease to have any relation towards extreme values. Intervening values would exhibit every variety of relationship, from the maximum in the neighbourhood of the means to zero value towards the extremes. These properties of the surface explain why Q is not even approximately constant, for those numerous surfaces of statistical practice in which the regression is approximately constant, i.e. linear.

Plates XXX, I (a) and (b), XXXI, I (c) give photographs of the actual model surface. In Fig. I (a) we see the surface 'end on' to one set of cross-sections and we grasp readily the skewness of the cross-sections and its increasing value as we pass from the central section. Fig. I (b) indicates by the verticals the regression line of means; these verticals have their feet on the regression line and the eye sees at once by their closer and closer approximation to each other, the deviation from linearity. Fig. I (c) gives a diagonal view of the surface of constant association, and underneath it Fig. II (d) has been placed a model of the Gaussian surface* of constant regression for comparison; the angle of deviation was taken roughly about 60°, to give a surface of correlation 0·5. Except for the regression line and skew-sections in Fig. I, the eye does not distinguish very readily in these photographs of Figs. I and II, the fundamental differences of the two surfaces which appear so markedly in the isoplethes of Diagram II.

* On the Brill system of interlaced sections.

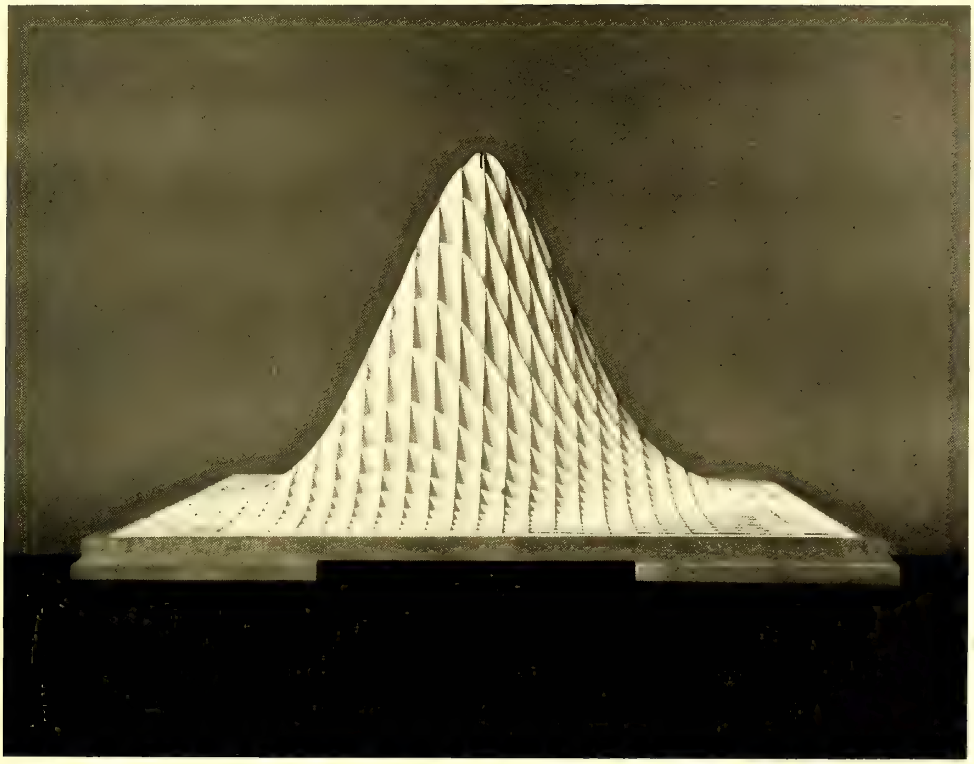


FIG. I (a). Surface of Constant Association.

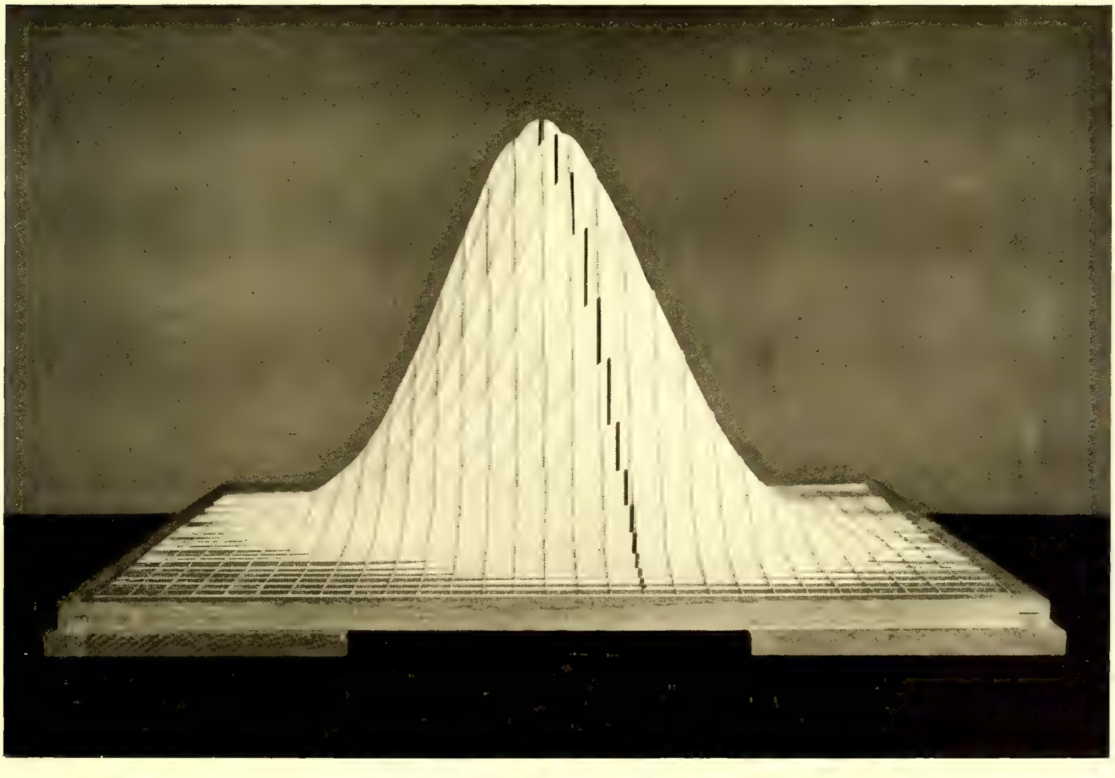


FIG. I (b). Surface of Constant Association.

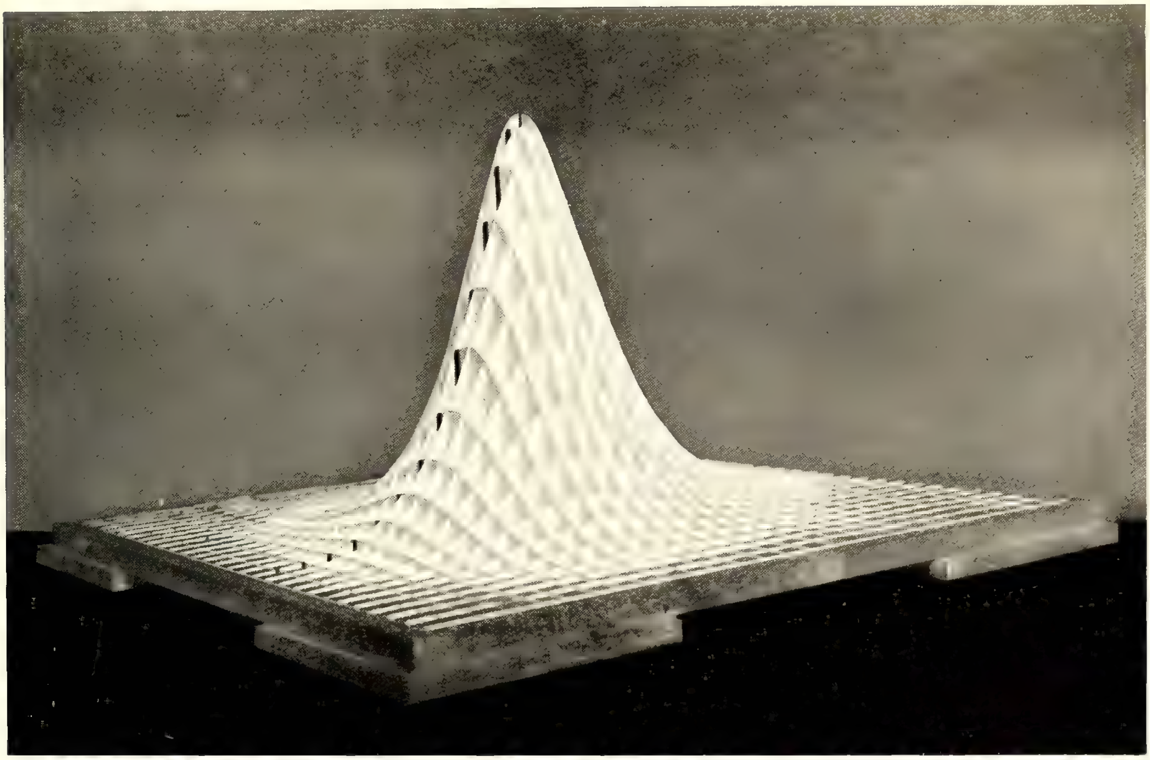


FIG. I (c). Surface of Constant Association.

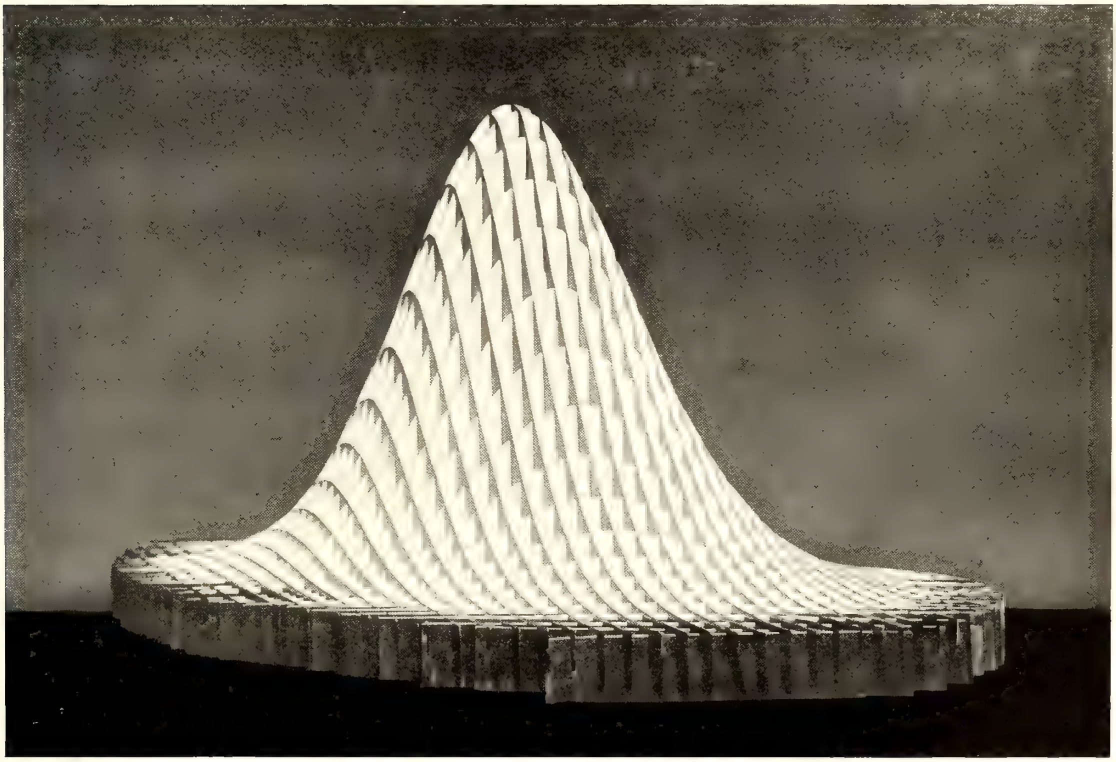
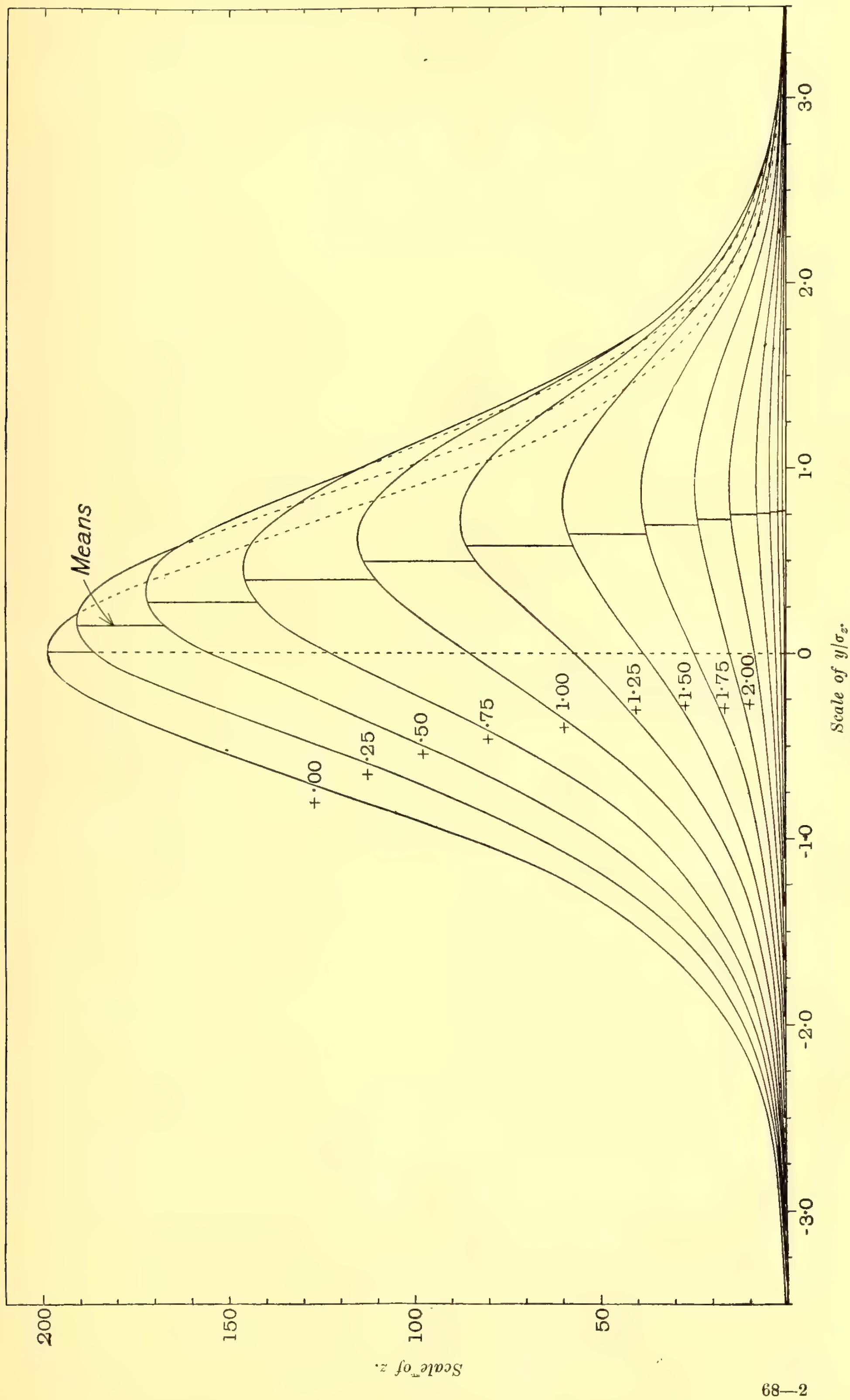


FIG. II (d). Brill type model of Gaussian surface.

DIAGRAM I. Surface of Constant Association, $Q=0{\cdot}6$.



Showing skew-sections at intervals of $x=0{\cdot}25\sigma_x$.

The equation to the surface of $Q=0{\cdot}6$ with normal marginal frequencies, for a total population of 1000, is:

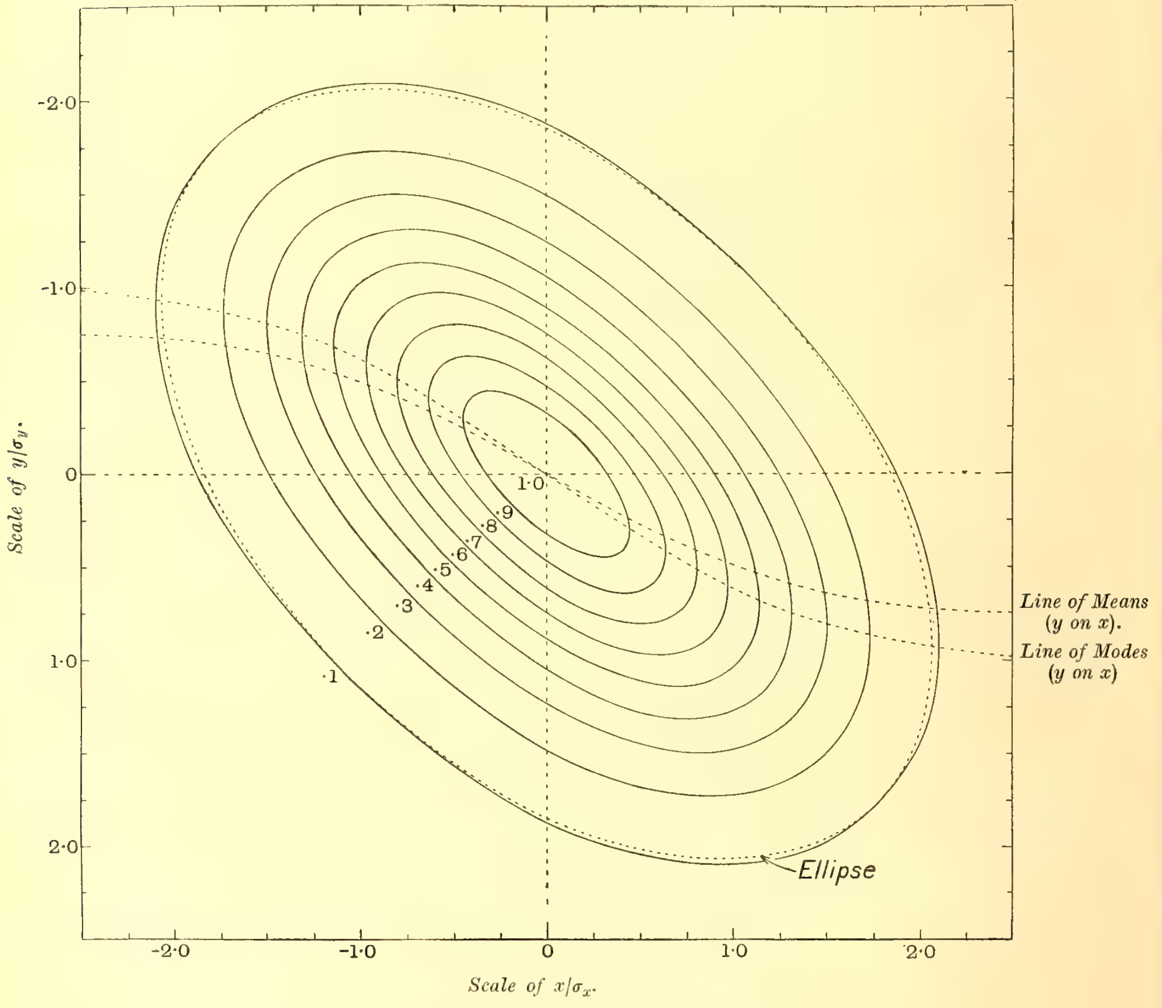
$$z=\frac{4000\dfrac{dp}{dx}\dfrac{dq}{dy}(1+3(p+q)-6pq)}{\{(1+3(p+q))^2-48pq\}^{\frac{3}{2}}},$$

where $p=\frac{1}{2}(1+\alpha)$ for x/σ_x,

$q=\frac{1}{2}(1+\alpha')$ for y/σ_y,

α having the usual significance attributed to it in Tables of the Probability Integral, e.g. Sheppard's.

DIAGRAM II. Isoplethes of Surface of Constant Association, $Q=0{\cdot}6$.



The isoplethes are drawn for each tenth of the maximum value of z.

The following short table will be sufficient to indicate the nature of the cross-sections to any-one wishing to pursue the matter further.

z-ordinates of y-arrays for given special values of x/σ_x.

Values of y/σ_y.	Array at $x/\sigma_x =$							
	0·0	*0·5*	*1·0*	*1·5*	*2·0*	*2·5*	*3·0*	*3·5*
3·5	0·223	0·332	0·388	0·314	0·165	0·059	0·015	0·003
3·0	1·135	1·688	1·970	1·589	0·834	0·297	0·077	0·015
2·5	4·536	6·736	7·811	6·223	3·230	1·145	0·297	0·059
2·0	14·881	21·381	24·195	18·497	9·269	3·230	0·834	0·165
1·5	38·080	55·053	57·904	39·890	18·497	6·223	1·589	0·314
1·0	85·096	113·670	100·499	57·904	24·195	7·811	1·970	0·388
0·5	155·888	168·289	113·670	55·053	21·381	6·736	1·688	0·332
0·0	198·944	155·888	85·096	38·080	14·881	4·536	1·135	0·223
−0·5	155·888	97·562	49·567	22·036	8·414	2·644	0·662	0·130
−1·0	85·096	49·567	25·217	11·369	4·379	1·381	0·346	0·068
−1·5	38·080	22·036	11·369	5·186	2·010	0·635	0·159	0·031
−2·0	14·881	8·414	4·379	2·010	0·781	0·247	0·062	0·012
−2·5	4·536	2·644	1·381	0·635	0·247	0·078	0·020	0·004
−3·0	1·135	0·662	0·346	0·159	0·062	0·020	0·005	0·001
−3·5	0·223	0·130	0·068	0·031	0·012	0·004	0·001	0·000
Means	0·000	0·268	0·494	0·646	0·723	0·757	0·766	0·769

III.

Studies in the Meaning and Relationships of Birth and Death Rates. I. The Relationship between "corrected" Death-rates and Life Table Death-rates, by JOHN BROWNLEE, M.D., D.Sc. Journal of Hygiene, Vol. XIII. No. 2, pp. 178—190.

There are three "death-rates" used by those who deal with the statistics of public health etc.; the "crude" death-rate found by dividing the total number of deaths in a district in one year by the total number of inhabitants; the "corrected" death-rate found by applying the death-rates for age and sex to a standard population and calculating the rate from the figures so found; and the "life table" death-rate which is the rate that would be found by working out a complete stationary population from the death-rates for each age and calculating the ratio of the total deaths among the assumed stationary population to the total stationary population. This result is simply the reciprocal of the "expectation of life" at birth. Dr Brownlee thinks that this last measure is the most satisfactory death-rate and his paper is an attempt to reach approximate values for it from the corrected death-rates. Although Dr Brownlee does not set them out in that way these approximations appear to be based on the use of an imaginary stationary population or populations and require, we think, to be tested more extensively than has yet been done before they are used for any practical conclusions.

It appears to us that the use of the "expectation of life" at birth or of a single death-rate for a population as a whole cannot, however it is calculated, gauge completely the mortality of that population, and cannot therefore form an entirely satisfactory basis for comparative purposes: two populations may show the same death-rate or expectation of life and yet the mortality of one of them may, for instance, be heavy only in the first five years of life, while the other is light for the first five years but heavy at later ages. There must therefore be limitations in the use of any single figure, and if more than one figure is used we do not see why the death-rates at various ages should not be employed as they stand. They are easier to interpret and no more difficult to deal with than the expectation of life at a series of ages,—an alternative measure implied by Dr Brownlee.

In the paper before us Dr Brownlee mentions that further work is to follow, and possibly it will show how he intends to overcome these difficulties. W. P. E.

IV. A Case of Familial Malformation in a Fowl's Head.

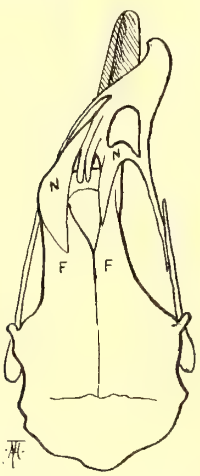
By HUBERT M. TURNBULL, M.D., Director of the Pathological Institute of the London Hospital.

S.D. 2105, 1912. *Fowl's head received from Mr Roxburgh, 10th October*, 1912.

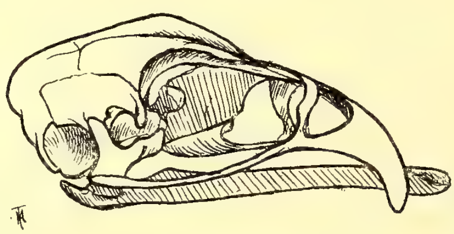
The breeder says that another chicken of the same brood has a similar deformity. Fowl examined turned out to be the heaviest of the whole brood, in spite of its crooked beak.

External appearance. Head and neck of domestic fowl, ?cock. Wattles are red and symmetrical. Flat, tongue-like, red comb. Eyes are of equal size. Above the nostril on the left side there is a semilunar cartilage 1·5 cms. long and 0·5 cm. from above down; nostril is 2·1 cms. from the anterior canthus. Right nostril is 1·3 from anterior canthus; semilunar cartilage is 0·5 by 0·3 cm. Nostrils on both sides are 2 cms. from the angle of the mouth. Horny beak starts on the left side 2·3 cms. from the angle of the mouth, on the right side 1 cm. from the angle. As a result the superior mandible on the left side is convex and the upper beak is twisted over to the right so that its point lies 0·2 cm. to the right of the right border of the lower beak, at a point 0·7 cm. from the end of the latter. The upper jaw and the beak on the right side form a sharp concavity. The anterior border of the orbit on the left side appears much farther back than on the right. Tongue is quite symmetrical. On the roof of the mouth the palate is divided into three divisions, one side having been cut longitudinally. The anterior part of the palate corresponding to the beak is not divided. This portion is deflected with the beak to the right side.

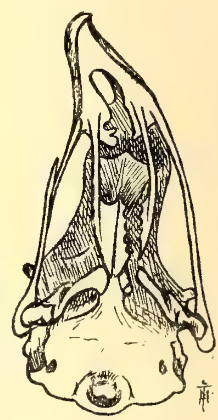
The chief alteration in the soft parts is the small size of the right nostril and its operculum. A longitudinal incision had been made, in killing the fowl, through the palate to the left of the middle line.



I. View of skull from above. Traced from a photograph.
N = Nasal bones.
F = Frontal bones.



III. View of skull from right side. Traced from a photograph.



II. View of skull from below. Traced from a photograph. Left palatal and pterygoid bones severed.

The head was simmered for five hours and the flesh removed with forceps. Partially cleaned in alcohol and chloroform—equal parts.

Skeleton. The incision in the palate has severed the left palatal and pterygoid bones close to their articulation with the vomer and rostrum. The right nostril measures 0·7 cm. long by 1 cm. high, the left 2 by 0·7 cm. The right frontal anterior to the posterior limit of the orbit is 3 cms. long and from a breadth posteriorly of 1 cm. tapers gradually; its anterior extremity appears to be of denser structure than that of the left frontal, and is devoid of a process mesial to the nasal bone. The left frontal is 2 cms. long and maintains a breadth of about 1 cm. This asymmetry throws the anterior part of the skull a little to the left. The superior surface of the right nasal bone measures 2·3 by 0·5 cms., the left 2·6 by 0·7 cms.; the lateral process of the right nasal does not share in the hypoplasia. The left nasal bone is thin and translucent, the right is thick and opaque. The right jugular and quadrato-jugular measure 4 cms. as against 4·2 cms. on the left. The right superior maxilla appears to be smaller than the left. The shortening of these bones on the right side bends the fore part of the face and beak sharply to the right. The other bones are symmetrical.

The narrow right frontal may be the result of hypoplasia of the right olfactory bulb, in consequence of which the soft and bony parts of the right nose may have failed to develop. This theory, however, does not give a satisfactory explanation of the elongation of the right frontal, the absence of its process mesial to the nasal bone, and the sclerosis of the right nasal bone. A more satisfactory explanation of the deformity is that it is the result of a hypoplasia of the right nasal bone, the elongation of the right frontal illustrating the interdependence of growing structures. The sclerosis of the right nasal bone, and apparently also of the external frontal process with which it articulates, suggest that the arrest in development has been caused by inflammation or trauma. The absence of the mesial process of the frontal may be due to involvement of the tissue in which it is formed, in such inflammation or trauma.

V. Note on a Ready Reckoner for the Obstetrician.

EDITORIAL.

On p. 503 of this volume of *Biometrika* two formulae are given by Dr de Souza, namely: (c') for determining the most probable value of the obstetric conjugate in terms of the known values of the diagonal conjugate and of the antero-posterior diameter, and further (i') giving the most probable value of the obstetric conjugate as deduced from the known values of the diagonal conjugate and of the height of the right hip-bone. It was proposed to include in Dr de Souza's paper two abacs or ready reckoners for the use of the obstetrician, so that after measuring on the living subject the diagonal conjugate and either the antero-posterior diameter or the height of the right hip-bone, he could read off from them the probable value of the obstetric conjugate without any calculation. These abacs were prepared by Miss H Gertrude Jones, but owing to certain difficulties of engraving, it was not found possible to delay the printing off of Dr de Souza's paper till they were completed. The two abacs have only reached us, in but moderately successful reproduction just as we go to press, and are included in this final note.

The method of using them is extremely simple. For example, in the case of Abac I, Plate XXXII, suppose the observed antero-posterior diameter to be 11·7 cms. and the diagonal conjugate to be 13·4 cms. we move upwards between the vertical lines at 11·6 and 11·8 on the antero-posterior diameter scale, until we meet the horizontal line at 13·4 cm. on the diagonal conjugate scale; we then pass off on the nearest sloping line to the scale of the obstetric

conjugate: the required sloping line lies between the 11·4 and 11·5 obstetric conjugate lines or the probable value of the obstetric conjugate is 11·45 cms. The actual value found from Dr de Souza's formula (c') is 11·43 cms. If the vertical antero-posterior diameter line and the horizontal diagonal conjugate line intersect in any case below *DD*, there is possibility of difficulty in labour.

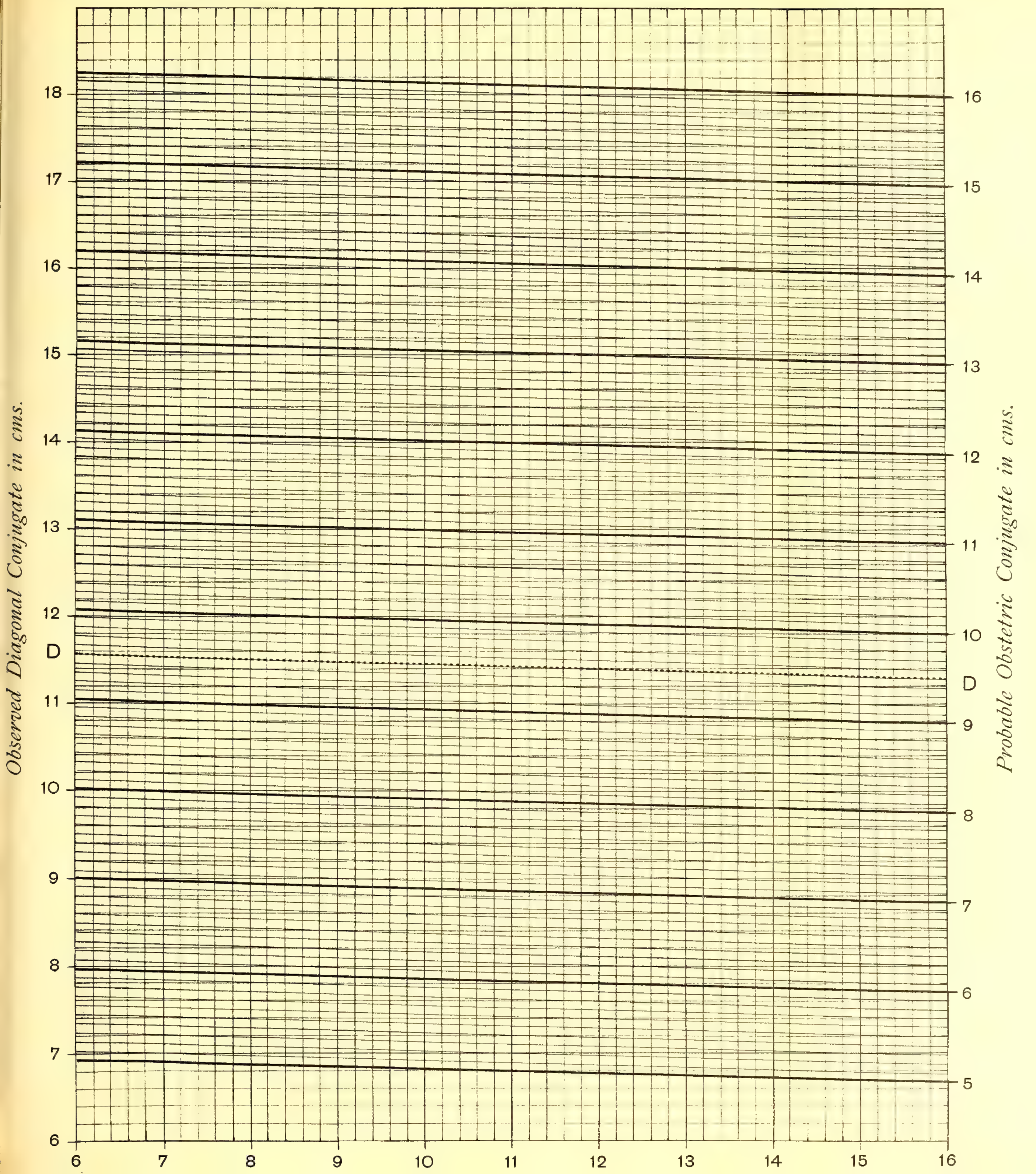
Again, if the diagonal conjugate be 13·4 and the height of the right hip-bone 20·7 cms., we enter by the vertical line at 20·7 on the right hip-bone scale of Abac II, Plate XXXIII, and the horizontal line on the diagonal conjugate scale at 13·4 and find that they meet on the sloping obstetric conjugate line of 11·45 cms. The actual value as found from formula (i'') is 11·44. The line of possible difficulty in labour is indicated by *DD*. The use of a lens may occasionally be of assistance, but most obstetricians will find themselves able to read the value of the probable obstetric conjugate to at least a millimetre and with a probable error of 3 mm., this is ample. Descriptions of the methods of measuring the required lengths on the living will be found in Dr de Souza's paper.

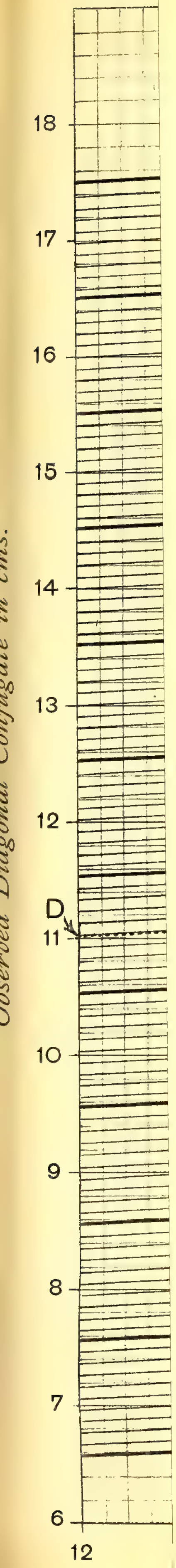
ADDENDUM. *Deformity known as Split-Foot, Biometrika*, Vol. IX. p. 385.

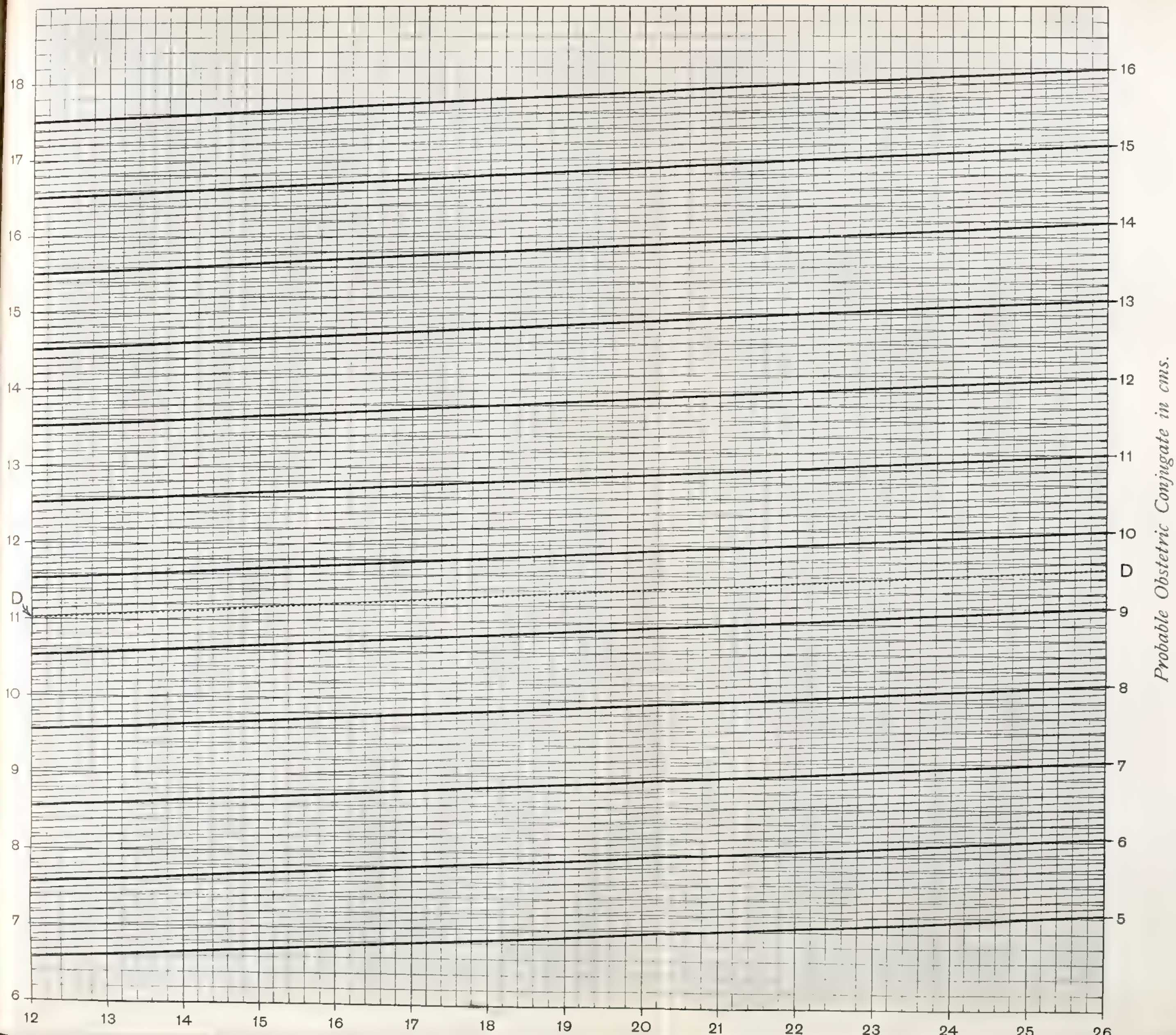
Annie Eliza S. (Mrs F.) IV. 9, has given birth a few days ago to a female child, quite normal in every respect, October 22, 1913.

CAMBRIDGE: PRINTED BY JOHN CLAY, M.A. AT THE UNIVERSITY PRESS

 OBSTETRICIAN'S ABAC I.




Observed Diagonal Conjugate in cms.
18
17
16
15
14
13
12
D
11
10
9
8
7
6
12



18
17
16
15
14
13
12
D
11
10
9
8
7
6
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
16
15
14
13
12
11
10
D
9
8
7
6
5
Probable Obstetric Conjugate in cms.

CONTENTS

The publication of a paper in *Biometrika* marks that in the Editor's opinion it contains either in method or material something of interest to biometricians. But the Editor desires it to be distinctly understood that such publication does not mark assent to the arguments used or to the conclusions drawn in the paper.

Biometrika appears about four times a year. A volume containing about 500 pages, with plates and tables, is issued annually.

Papers for publication and books and offprints for notice should be sent to Professor KARL PEARSON, University College, London. It is very desirable that a copy of all measurements made, not necessarily for publication, should accompany each manuscript. In all cases the papers themselves should contain not only the calculated constants, but the distributions from which they have been deduced. Diagrams and drawings should be sent in a state suitable for direct photographic reproduction, and if on decimal paper it should be *blue* ruled, and the lettering only pencilled.

Papers will be accepted in German, French or Italian. In the first case the manuscript should be in Roman not German characters.

Contributors receive 25 copies of their papers free. Fifty additional copies may be had on payment of 7/- per sheet of eight pages, or part of a sheet of eight pages, with an extra charge for Plates; these should be ordered when the final proof is returned.

The subscription price, payable in advance, is 30*s.* *net* per volume (post free); single numbers 10*s.* *net.* Volumes I, II, III, IV, V, VI, VII, VIII and IX (1902—13) complete, 30*s.* *net* per volume. Bound in Buckram 34/6 *net* per volume. Index to Volumes I to V, 2*s.* *net.* Subscriptions may be sent to C. F. Clay, Cambridge University Press, Fetter Lane, London, E.C., either direct or through any bookseller, and communications respecting advertisements should also be addressed to C. F. Clay.

Till further notice, new subscribers to *Biometrika* may obtain Vols. I—IX together for £10 *net*—or bound in Buckram for £12 *net.*

CAMBRIDGE: PRINTED BY JOHN CLAY, M.A. AT THE UNIVERSITY PRESS.